DATA EVALUATION AND METHODS RESEARCH

Series 2
Number 77

On the Mathematics of Competing Risks

This report reviews the sources and origins of competing risk theory and presents the probability theory of competing risks and statistical estimation techniques and tests of hypotheses.

DHEW Publication No. (PHS) 79-1351

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Public Health Service
Office of the Assistant Secretary for Health
National Center for Health Statistics
Hyattsville, Md. January 1979

Library of Congress Cataloging in Publication Data

Birnbaum, Z. William.
On the mathematics of competing risks.

(Vital and health statistics: Series 2, Data evaluation and methods research; no. 77) (DHEW publication; (PHS) 79-1351)
Bibliography: p. 56
1. Competing risks. 2. Estimation theory. 3. Statistical hypothesis testing. I. Title. II. Series: United States. National Center for Health Statistics. Vital and health statistics: Series 2, Data evaluation and methods research; no. 77. III. Series: United States. Dept. of Health, Education, and Welfare. DHEW publication; (PHS) 79-1351.
RA409.U45 no. 77 [QA273] 312'.07'23s [362.1'01'82]
ISBN 0-8406-0138-7 78-15122

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402

NATIONAL CENTER FOR HEALTH STATISTICS

OFFICE OF THE ASSOCIATE DIRECTOR FOR MATHEMATICAL STATISTICS

Vital and Health Statistics-Series 2-No. 77

DHEW Publication No. (PHS) 79-1351
Library of Congress Catalog Card Number 78-15122

FOREWORD

In view of the many potential applications of the statistical theory of competing risks in the analysis of health and vital statistics, the National Center for Health Statistics was responsive to Dr. Birnbaum's proposal to prepare an introductory report on the mathematics of competing risks which would clarify the concepts and unify the theory common to the several disciplines in which competing risk models have been applied. As a result, we have this excellent report.

It has become Center policy to submit all reports in this publication series to peer review. In this regard, we should like to recognize the contributions of Professor Anthony J. Quinzi who served as technical reviewer of this report. Dr. Birnbaum joins me in thanking Dr. Quinzi for his careful review and for his many helpful suggestions and comments.

Monroe G. Sirken
Associate Director for
Mathematical Statistics

CONTENTS

LIST OF FIGURES

TABLE

ON THE MATHEMATICS OF COMPETING RISKS

Z. William Birnbaum, University of Washington, Seattle

1. INTRODUCTION

1.1. SOURCES AND ORIGINS OF COMPETING RISK THEORY

1.1.1. Work of Daniel Bernoulli

Toward the middle of the 18th century, the question of mandatory vaccination against smallpox was widely discussed and attempts were made to evaluate its possible effects. In his "Mémoir" published in 1776 (figure 1), Daniel Bernoulli[1] gave this question the following specific formulation:

Available life tables reflect the mortality of the population for which they were calculated, taking into account all causes of death including smallpox. How would these life tables change if, because of mandatory vaccination, deaths from smallpox were entirely eliminated?

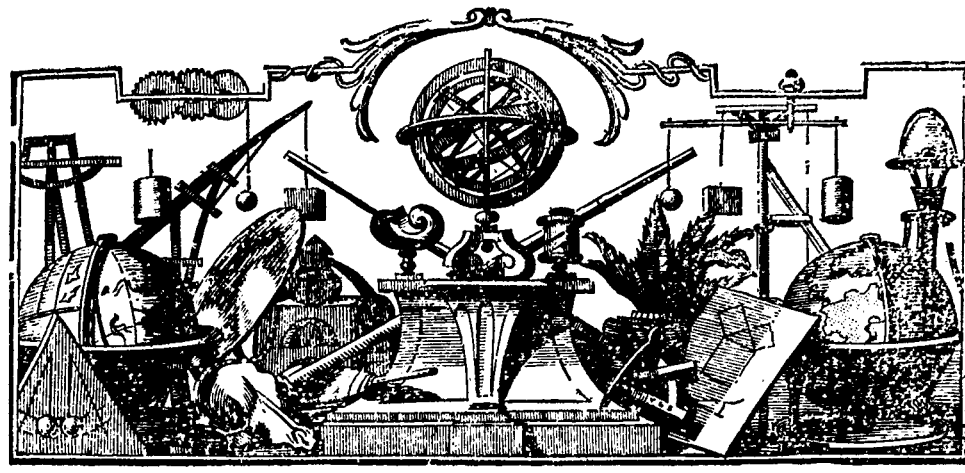
As empirical sources of information, Bernoulli considered existing life tables, such as those computed in 1693 by the astronomer E. Halley[2] from the records of the city of Breslau in Germany. These tables were based on data which reported age at death for each individual; in addition, it may have been possible to determine whether death was a result of smallpox or of other causes. Therefore, each individual may be considered as exposed to two "risks": death from smallpox and death from other causes. These risks competed for his life in the sense that the risk which materialized first determined his life length. Thus for an individual who died of smallpox there was no way of telling how long he would have lived had smallpox been eliminated—and it was just this missing information that Bernoulli needed to answer his question.

The argument devised by Bernoulli is ingenious and rather simple. It is also open to serious criticisms as to assumptions made explicitly or by implication. These criticisms already point at most of the difficulties encountered later in dealing with competing risks. We shall outline Bernoulli's argument,[1] following closely the somewhat naive reasoning of his Mémoir and only modifying his notation.

Consider a population of l_0 individuals of age 0, and let l_x denote the number of those among the l_0 who are still alive at age $x > 0$. Assume l_x known for $x \geqslant 0$, for example from a life table. Clearly, l_x is a decreasing function of x.

Let s_x be the number of those of the l_0 who survive to age $x > 0$ and who have been infected with and survived smallpox. They are immune from smallpox for the rest of their lives. Similarly, let t_x be the number of those surviving to age $x > 0$ who did not have smallpox, hence may still contract this disease. Clearly,

$$l_x = s_x + t_x .$$



MÉMOIRES
DE
MATHÉMATIQUE
ET
DE PHYSIQUE,
TIRÉS DES REGISTRES
de l'Académie Royale des Sciences,
De l'Année M. DCCLX.

ESSAI D'UNE NOUVELLE ANALYSE
De la mortalité cauſée par la petite Vérole, & des avantages de l'Inoculation pour la prévenir.

Par M. DANIEL BERNOULLI.

INTRODUCTION APOLOGÉTIQUE.*

CEUX qui ont ſenti tout l'avantage de l'Inoculation, ont imaginé différentes façons de repréſenter cet avantage, qui, quoique revenant au même, ne laiſſent pas de faire une

* Cette Introduction n'a été faite que long-temps après le Mémoire, étant du 16 Avril 1765.

Figure 1. Introductory page of Daniel Bernouilli's monograph on his theory of competing risks.

Assume that in a unit of time (e.g., a year) smallpox attacks 1 in n individuals alive, and causes the death of 1 in m of those attacked. Bernoulli assumed m and n to be known (he estimated both to be about 8), and independent of x.

The number of those dying of all causes during a time interval $(x, x + dx)$ is

$$l_x - l_{x+dx} = -dl_x .$$

Among them are

$$\frac{t_x\, dx}{nm} = \text{number of those who get smallpox and die of the disease} \tag{1}$$

and

$$-dl_x - \frac{t_x\, dx}{nm} = \text{number of those who die of other causes.} \tag{2}$$

The number t_x decreases during the time interval $(x, x + dx)$ by the number of those among the t_x who get smallpox and by the number of those among the t_x who die of causes other than smallpox, so that

$$t_x - t_{x+dx} = -dt_x = t_x \frac{dx}{n} + \frac{t_x}{l_x}\left(-dl_x - \frac{t_x\, dx}{nm}\right).$$

This can be rewritten

$$n\, d\left[\log\left(m\frac{l_x}{t_x} - 1\right)\right] = dx,$$

which yields by integration

$$t_x = \frac{ml_x}{1 + \exp\left[(x + c)/n\right]},$$

and since $t_0 = l_0$, one has

$$\exp\,(c/n) = m - 1,$$

so that

$$t_x = \frac{ml_x}{1 + (m - 1)\exp\,(x/n)}. \tag{3}$$

Consider now the life table obtained from the l_x by eliminating smallpox as a cause of death, and denote z_x = number of those surviving to age $x > 0$ when smallpox is eliminated, with $z_0 = l_0$. In the presence of smallpox, the numbers of those dying of each of the two causes during a time interval $(x, x + dx)$ were given by equations (1) and (2). Without smallpox only equation (2) applies, but at x there are z_x alive instead of l_x, hence the number of those dying in a time interval $(x, x + dx)$ without smallpox is

$$-dz_x = -\frac{z_x}{l_x}\left(dl_x - \frac{t_x\,dx}{nm}\right).$$

Substituting t_x from equation (3), one obtains

$$\frac{dz_x}{z_x} = \frac{dl_x}{l_x} + \frac{dx}{n[1 + (m-1)\exp(x/n)]},$$

a differential equation with the solution

$$\frac{z_x}{l_x} = \frac{m}{m - 1 + \exp(-x/n)}. \tag{4}$$

This ratio is clearly greater than 1 for $x > 0$, and it describes the improvement of the life table z_x over the original life table l_x. Asymptotically, one has for $x \to \infty$

$$\frac{z_x}{l_x} \to \frac{m}{m-1}.$$

Using this mathematical model with the values $m = n = 8$ for which he claimed some empirical justification, Daniel Bernoulli calculated his Tables I and II reproduced as they appear in his original paper (figure 2). Column headings were added for present use.

In Table I the first and second columns (x and l_x) are taken from Halley's life table; the third column contains values of t_x calculated from equation (3) with $m = n = 8$, and the fourth column lists the values $s_x = l_x - t_x$. The number of those contracting smallpox at age x appears in column 5 and is computed as follows: On the intuitive assumption that the number t_x decreases uniformly during a year, consider $(1/2)(t_x + t_{x+1})$ as the number of those exposed to the risk of contracting smallpox at age x, and enter under column 5 the estimate $(1/2)(t_x + t_{x+1})/8$. Dividing this number by $m = 8$, one obtains the entry in column 6, an estimate of the number of those dying of smallpox in a year. Column 7 accumulates the numbers of smallpox deaths from column 6, and column 8 accumulates the deaths from other causes.

In Table II, the first two columns contain Halley's life table reflecting mortality from smallpox and from all other causes; the third column contains values of z_x computed according to equation (4). The fourth column lists the "gains" $z_x - l_x$.

A number of criticisms were raised against certain of Bernoulli's assumptions even before publication of his paper—some he ascribed to an "eminent mathematician," most likely d'Alembert, and he

MÉMOIRES DE L'ACADÉMIE ROYALE DES SCIENCES.

TABLE I.

(1) x	(2) l_x	(3) t_x	(4) s_x	(5)	(6)	(7)	(8)
AGES par années.	Survivans selon M. Halley.	N'ayant pas eu la pet. vérole.	Ayant eu la pet. vérol.	Prenant la pet. vérole pendant ch. année.	MORTS de la pet. vérole pendant chaq. ann.	SOMME des morts de la pet. vérole.	MORTS par d'autres maladies pend. chaq. année.
0	1300	1300	0				
1	1000	896	104	137	17,1	17,1	283
2	855	685	170	99	12,4	29,5	133
3	798	571	227	78	9,7	39,2	47
4	760	485	275	66	8,3	47,5	30
5	732	416	316	56	7,0	54,5	21
6	710	359	351	48	6,0	60,5	16
7	692	311	381	42	5,2	65,7	12,8
8	680	272	408	36	4,5	70,2	7,5
9	670	237	433	32	4,0	74,2	6
10	661	208	453	28	3,5	77,7	5,5
11	653	182	471	24,4	3,0	80,7	5
12	646	160	486	21,4	2,7	83,4	4,3
13	640	140	500	18,7	2,3	85,7	3,7
14	634	123	511	16,6	2,1	87,8	3,9
15	628	108	520	14,4	1,8	89,6	4,2
16	622	94	528	12,6	1,6	91,2	4,4
17	616	83	533	11,0	1,4	92,6	4,6
18	610	72	538	9,7	1,2	93,8	4,8
19	604	63	541	8,4	1,0	94,8	5
20	598	56	542	7,4	0,9	95,7	5,1
21	592	48,5	543	6,5	0,8	96,5	5,2
22	586	42,5	543	5,6	0,7	97,2	5,3
23	579	37	542	5,0	0,6	97,8	6,4
24	572	32,4	540	4,4	0,5	98,3	6,5

Figure 2. Daniel Bernouilli's original life tables. (Column headings were added for present use.)

discussed them at some length. The more obvious ones (e.g., the assumption that the probability of contracting smallpox as well as the probability of dying of smallpox are independent of age), could be replaced by more flexible and realistic assumptions without much difficulty. Other assumptions, such as the use of $(1/2)(t_x + t_{x+1})$ in estimating the number of those attacked by smallpox for column 5 in Table I, have been reappearing throughout the centuries, and we will have to deal with them later on.

MÉMOIRES DE L'ACADÉMIE ROYALE DES SCIENCES.

TABLE II.

(1) x	(2) l_x	(3) z_x	(4)	(1) x	(2) l_x	(3) z_x	(4)
AGES par années.	État naturel & variolique.	ÉTAT non-varioliq.	Différ. ou gains.	AGES par années.	État naturel & variolique.	ÉTAT non-varioliq.	Différ. ou gains.
0	1300	1300	0	13	640	741,1	74,1
1	1000	1017,1	17,1	14	634	709,7	75,7
2	855	881,8	26,8	15	628	705,0	77,0
3	798	833,3	35,3	16	622	700,1	78,1
4	760	802,0	42,0	17	616	695,0	79,0
5	732	779,8	47,8	18	610	689,6	79,6
6	710	762,8	52,8	19	604	684,0	80,0
7	692	749,1	57,2	20	598	678,2	80,2
8	680	740,9	60,9	21	592	672,3	80,3
9	670	734,4	64,4	22	586	666,3	80,3
10	661	728,4	67,4	23	579	659,0	80,0
11	653	722,9	69,9	24	572	651,7	79,7
12	646	718,2	72,2	25	565	644,3	79,3

Cette Table fait voir d'un coup d'œil, combien ſur 1300 enfans, ſuppoſés nés en même temps, il en reſteroit de vivans d'année en année juſqu'à l'âge de vingt-cinq ans, en les ſuppoſant tous ſujets à la petite vérole; & combien il en reſteroit s'ils étoient tous exempts de cette maladie, avec la comparaiſon & la différence des deux états.

Figure 2. Daniel Bernouilli's original life tables. (Column headings were added for present use.)—Con.

Two features of Bernoulli's approach are worth noting. One is the deterministic interpretation of all processes: The quantities l_x, s_x, t_x, z_x are functions of age x. A probabilistic approach could hardly have been expected at that time. The other feature is the nonparametric treatment: Bernoulli made no assumptions about the functional form of l_x and, consequently, of the other functions of age. It is only with the arrival of the life-table models introduced by Gompertz in 1825[3] and Makeham in 1860[4] that specific functional forms of l_x were considered.

1.1.2. Actuarial Problems

More than one hundred years after Bernoulli's Mémoir,[1] W. M. Makeham in 1874[5] developed a theory of "multiple decremental forces," and thereby gave a systematic foundation to the actuarial treatment of competing risk problems. At Makeham's time, the theory of probability had already been developed and actuaries were formulating their statements in probabilistic terms. Their earlier work was what, in discussing Bernoulli's approach, we referred to as nonparametric: They started out with life tables obtained from experience without ascribing to them a specific functional form and estimated all the probabilities and actuarial values required. In 1860, Makeham,[4] expanding on earlier work by B. Gompertz in 1825,[3] introduced the important parametric model known as the Gompertz-Makeham law of mortality which appeared to be particularly suitable as an approximation to many empirical life tables.

Actuarial theory has traditionally used both approaches: nonparametric as well as applying the Gompertz-Makeham law. In classical works such as E. F. Spurgeon's in 1932 (pp. 223-320[6]) both treatments are discussed in great detail, and it is pointed out that, if a life table follows Makeham's law, the mathematical arguments become more elegant and manageable.

The competing risks problems considered by actuaries were usually of a different kind and simpler than the question asked by Bernoulli. Such problems arose mostly when insurance policies had to be written on several lives, e.g., a policy for husband and wife that assured the payment of a benefit (such as a lump sum, or life annuity) after the death of one of them, to the surviving spouse. Here the premium to be charged is determined by the probability distribution of the time to the earlier death, hence could be computed if the joint probability distribution of both lives were available. Since life tables for pairs of lives, which would yield such joint probability distributions, are difficult to obtain, actuarial discussions of competing risks proceed on the assumption that the lives involved are independent in the sense of probability theory. This assumption has been recognized as unrealistic under many circumstances. For example, two people living together are likely to be exposed to the same environment, including a number of hazards, which affects their lives in a similar manner. Nevertheless, little, if anything, has been done in actuarial theory to account for this kind of stochastic dependence.

When life tables are used to estimate various probabilities, without the benefit of parametric models such as the Gompertz-Makeham law, actuaries have been facing the problem which Bernoulli had to deal with in calculating column 5 of Table I: If life-table-type data are available only at endpoints of fixed time intervals, how does one use them when they appear in the denominator of a formula for a probability? The following procedure, quite analogous to Bernoulli's, was recommended by Spurgeon in 1932 (p. 382[6]). Denoting by E_x the number of individuals alive at age x, and by θ_x the number of the E_x who die between ages x and $x + 1$, he states:

"If a number of the E_x persons aged x, say w_x, withdrew from observation during the year, it would not be known whether they survived to age $x + 1$ or died between the date of withdrawal and the attainment of that age.... Assuming that the withdrawals were equally distributed throughout the year..., the rate of mortality will be obtained by taking

$$q_x = \frac{\theta_x}{E_x - (1/2)w_x}." \tag{5}$$

1.1.3. Clinical Experiments

A large number of clinical experimental studies have been carried out in a manner which, possibly somewhat oversimplified, can be described by the following scheme.

Included in an experiment are N subjects (patients, animals). At the beginning of the experiment each of them is exposed to a treatment that is conjectured to be somehow related to an event referred to as "failure." For each individual subject, the experiment can terminate in one of two ways: Either failure is observed or the individual is lost from observation before failure occurs. Observable is the time W from the beginning of the experiment to its termination. This time W is clearly the smaller of two quantities:

T = time to failure
L = time to loss from observation.

Typical experiments of this kind deal with the effect of potentially carcinogenic agents. Experimental animals are treated with such an agent and some develop cancer; others are lost from observation either by death from other causes or by developing conditions that would invalidate the experiment. An aim of the experiment is to gather information on the time from the beginning of the treatment to the onset of cancer. This information could then be used, e.g., for comparisons with similar data for animals who have not been exposed to the agent under study; or it could be used to assess the effect of age or sex or some other factors on the time to the occurrence of cancer. The crucial difficulty in using data obtained from such experiments is that one does not know the time to the occurrence of cancer in those experimental subjects whose observation is terminated by the competing risk, i.e., loss from observation. Many studies overlooked this difficulty and arrived at conclusions that could be shown to have no validity (for an example, see the 1939 article by Bernstein, Birnbaum, and Achs[7]).

Another class of experimental studies in which loss from observation and, possibly, other competing risks make it difficult to evaluate the results, consists of experiments in which a treatment is used to delay the occurrence of some event. Research on various prophylactic treatments falls into this class, as do studies on the effectiveness of birth-control procedures.

Lack of familarity with even the classical actuarial theory has often affected the validity of clinical research. It was only in the mid-1950's that papers by statisticians such as Cornfield[8] began to remedy this situation by clearly pointing out the problem and by presenting actuarial techniques to those engaged in medical research.

A model that is substantially more complex than that of competing risks, and yet quite realistic, was introduced in 1951 by Fix and Neyman.[9] They considered the study of a population of healthy individuals who may leave that population either by becoming sick or by dying; but by becoming healthy, the sick may return to the population. Treating such health-sickness-death sequences as a stochastic process, they stimulated a succession of publications that go beyond the competing-risks concept. A survey of work in this direction may be found in Chiang's book[10] published in 1960 and in his report[11] published in 1974.

The stochastic process approach has been particularly useful in constructing models for human reproduction. First systematically presented in 1963 by Shepps and Perrin[12] and followed by a sequence of papers developing the mathematical theory and, as an alternative, proposing computer simulation (see Perrin's 1967 study[13]), these models have led to a wide range of approaches. In 1973, Shepps and Menken[14] surveyed these developments.

1.1.4. Competing Risks in Technology

If a technological system, such as an electronic network or a mechanical device, consists of n components $c_1, c_2, \ldots, c_n$ in series, and each component has a random life length, then the life of the entire system will end with the failure of the shortest lived component. If one agrees to say that

system failure is a result of the kth risk when c_k is the shortest lived component, then the n different risks are competing in the sense that one usually can observe only the life length of the system and the component whose failure coincides with the failure of the system; in most practical situations, however, one cannot tell how long the remaining $n - 1$ components would have lasted. And yet such knowledge is essential if one wishes to make some statements on the extent to which the life of the system can be prolonged or maintained by one of various patterns of replacing or servicing components. For example, components can be replaced at fixed time intervals; or there is a finite supply of spare components that are used to replace a working component immediately when it fails, so that system failure can be postponed until all spares have been used and then a component fails; or a particular component can be "beefed up" to increase its life length.

Such problems have become particularly important since the advent of systems with large numbers of components, such as computers, contemporary aircraft, or the total equipment involved in the successful launching of a space rocket. Such advanced systems gave rise to a new discipline, the mathematical theory of reliability which has been developed mainly within the last two decades, most of it in the United States and the Soviet Union. Fortunately, those actively involved in this development represented a wide range of backgrounds and combined technological insight, knowledge of actuarial concepts and ability to use mathematical abstraction, so that much of the learning-by-mistake phase mentioned in section 1.1.3 could be avoided.

1.1.5. Other Areas

The sources of competing-risk problems mentioned previously do not constitute an exhaustive list. They provide representative examples, but many others can be added. In the area of population dynamics, birth and immigration and death and emigration exemplify various modes of increment or of decrement of a population, and lead to the posing of problems in demography or in the study of biological populations.

1.2. COMMENTS ON THE PRESENT STATE OF THE THEORY

1.2.1. Duplication and Communication Gaps

In surveying the literature, one encounters instances in which researchers in some fields have been unaware of progress made in other areas, and either had to duplicate work already done or else did not use that work. Consequently, their conclusions were often weaker then they could have been, or sometimes were invalid. One of the main aims of this report is to focus on basic concepts common to many areas in which competing risks have been considered and to present the mathematical theory dealing with these concepts.

1.2.2. Probabilistic Versus Statistical Approach

In publications dealing with competing risks, much space has been devoted to the study of probabilistic models and their properties. Most of the traditional results are such that, when either a life table or some parameters of an underlying process are given, one can compute probabilities of various events, or expectations of times to failure or of other random variables, their variances and covariances, and so forth.

How to use empirical data to arrive at conclusions about the underlying probabilities was a question which for a long time was dealt with by rather primitive methods. Actuaries used smoothing (graduation) techniques to obtain life tables from raw mortality data, then estimated probabilities by expres-

sions such as equation (5), and their practices were largely accepted and used by workers in such fields as vital statistics and clinical research. Only recently has the corresponding statistical theory begun to develop, and with this development came, among others, a critical investigation of the role of independence of the "lives" involved, a systematic discussion of the question of what information can be extracted from empirical data (problems of identifiability), of the advantages and disadvantages of traditional and of some newly proposed estimates, and of problems that can be stated in terms of tests of hypotheses. Part of this report will be devoted to such developments in statistical theory and to the resulting practical techniques.

1.2.3. Choice of Topics for This Report

This presentation does not aim at completeness in any sense of the word. Our chief aim is to formulate the mathematical concepts and the theory of competing risks common to various fields of practical application. Having done that, we state a number of problems and describe techniques for dealing with them. There is a large literature on such problems, and subjective choices had to be made in selecting those which appeared typical, illustrative, and of practical interest.

In presenting details of mathematical derivations, too, some choices had to be made. As a rule, mathematical arguments are completely included when they are reasonably simple and representative of a method of approach. For derivations that require long and complicated arguments or the use of esoteric mathematical tools, the reader is most frequently referred to original sources.

In many cases, the theory available at present is inadequate, mainly because it requires assumptions that are not satisfied in real situations, such as the assumption of independence of random quantities that in practice are likely to be dependent. In other cases, the theory deals with statistical estimates that are biased or not consistent. In still other cases, such as in dealing with problems of identifiability (see section 2.6), the theory leads to the negative result that the data, as they are usually available, cannot yield the information one would like to obtain. In such situations, where there seem to be no satisfactory answers, we at least try to formulate the problems and point out the difficulties.

2. PROBABILITY THEORY OF COMPETING RISKS

2.1 LIFE DISTRIBUTIONS

In studying living organisms, one often considers their life length—the length of time from birth to death. Similarly, for technological devices, one may be interested in the length of time from the instant when the device was put into use to the instant when it failed. In either case one deals with a random variable capable of assuming only nonnegative values. Random variables of this kind are encountered in many other situations, such as incubation periods of infectious diseases, hospital stays, and the time an individual remains under observation or participates in an experiment.

In the material that follows, we shall use the terms "life" or "life length" or "time to failure" quite generally to mean a nonnegative random variable, i.e., a random variable X such that

$$P(X < 0) = 0.$$

When a value x of a life X is observed, we shall say that the individual failed (or died) at a time (or age) x, that is, that it was alive (functioning or "up") at all times t: $0 \leqslant t < x$, and that it is failed (or dead, or "down") at all times t: $x \leqslant t < \infty$.

The probability distribution function of a life X will be denoted by

$$F_X(x) = P(X \leqslant x),$$

the probability that failure occurs not later than at time x. It will simplify many arguments if we assume that X has a probability density, i.e., that there is a function $f_X(x)$ such that

$$f_X(x) = 0 \qquad \text{for } x < 0$$

$$f_X(x) \geqslant 0 \qquad \text{for } x \geqslant 0$$

$$\int_{-\infty}^{+\infty} f_X(x)\, dx = 1$$

and

$$F_X(x) = \int_0^x f_X(s)\, ds.$$

The function

$$\begin{aligned} \bar{F}_X(x) &= 1 - F_X(x) \\ &= P(X > x) \\ &= \int_x^{\infty} f_X(s)\, ds \end{aligned}$$

will be called the "survival function," and the function

$$\lambda_X(x) = -\frac{d}{dx} \log \bar{F}_X(x) = f_X(x)/\bar{F}_X(x) \tag{6}$$

is known as the "hazard function" for life X. The intuitive meaning of $\lambda_X(x)$ becomes clear when one writes the last equality as an equation between probability elements

$$\lambda_X(x)\, dx = \frac{f_X(x)\, dx}{\bar{F}_X(x)}$$

and interprets the right-hand expression as the conditional probability that an individual, having survived to age x, fails in the time interval $(x, x + dx)$.

The relationships

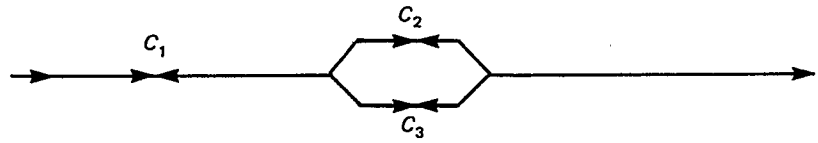
$$\bar{F}_X(x) = 1 - F_X(x)$$

$$= \exp\left[-\int_0^x \lambda_X(s)\,ds\right] \tag{7}$$

are easily verified and one observes that each of the functions $f_X(x)$, $F_X(x)$, $\bar{F}_X(x)$, $\lambda_X(x)$ determines all the others, hence any one of them describes the probability distribution of X.

2.2 NET (POTENTIAL) AND CRUDE (OBSERVABLE) LIVES

Example A.—We consider an electrical circuit described by the diagram



in which the only components subject to failure are the contacts C_1, C_2, C_3. We assume that each of these contacts deteriorates with time until failure occurs, and we denote the corresponding component lives by T_1, T_2, T_3, with probability densities

$$f_{T_1}(.), f_{T_2}(.), f_{T_3}(.).$$

We furthermore assume that T_1, T_2, T_3 are independent. Let W denote the time to failure of the circuit. Clearly the circuit fails when either C_1 fails, or C_2 and C_3 both fail, whichever occurs first. We shall refer to the failure of C_1 as "risk R_1," and to the failure of both C_2 and C_3 as "risk R_2," and define

$$\begin{aligned} X_1 &= \text{time to failure of } C_1 \\ &= T_1 \\ X_2 &= \text{time to failure of } C_1 \text{ and } C_2 \\ &= \max(T_2, T_3) \end{aligned}$$

Then the time to failure of the circuit is

$$W = \min(X_1, X_2)$$

$$= \min[T_1, \max(T_2, T_3)]$$

and its survival function can be computed as

$$\bar{F}_W(w) = P[T_1 > w, \max(T_2, T_3) > w]$$

$$= \int_w^\infty f_{T_1}(t) \left[\iint_{\max(t_2,t_3)>w} f_{T_2}(t_2) f_{T_3}(t_3)\, dt_3\, dt_2 \right] dt_1$$

$$= \int_w^\infty f_{T_1}(t_1) \int_w^\infty f_{T_2}(t_2) \int_{t_0}^{t_2} f_{T_3}(t_3)\, dt_3\, dt_2\, dt_1$$

$$+ \int_w^\infty f_{T_1}(t_1) \int_w^\infty f_{T_3}(t_3) \int_0^{t_3} f_{T_2}(t_2)\, dt_2\, dt_3\, dt_1.$$

Following an accepted terminology, we shall call X_1 and X_2 the "net" lives for risks R_1 and R_2, respectively. In practice, only the "actual" life W is observed, together with the fact that the circuit fails due to risk R_1 or risk R_2, so that the random variable observed is two-dimensional

$$(W, J)$$

where W is the actual life and $J = 1$ or 2, according to which of the two risks "caused" the failure.

The actually observed time to circuit failure due to risk R_1 is a random variable that has the conditional probability distribution of circuit life, given that failure is due to R_1. The random variable Y_1 with this probability distribution is referred to as the "crude" life for R_1. It has the survival function

$$\bar{F}_{Y_1}(y_1) = P(W > y_1 \mid J = 1)$$

$$= \frac{P(W > y_1, J = 1)}{P(J = 1)}$$

One similarly defines the "crude" life Y_2, due to R_2.

Example B.—A cohort of patients is under observation during a fixed period of time $0 \leqslant t \leqslant \tau$. For each individual, observation may end due to either of the two "risks": The individual develops a specified condition C, or the individual drops out of observation. We consider the random variables:

X_1 = time until condition C occurs
X_2 = time to dropping out from observation.

These two random variables will be referred to as the "net" lives for the first or second risk.

We assume that X_1 and X_2 are independent, and that X_1 has a probability density that increases with time according to the formula

$$f_{X_1}(x_1) = \begin{cases} 2x_1/\tau^2 & \text{for } 0 \leqslant x_1 \leqslant \tau \\ 0 & \text{elsewhere} \end{cases}$$

while X_2 has the uniform probability density

$$f_{X_2}(x_2) = \begin{cases} 1/\tau & \text{for } 0 \leqslant x_2 \leqslant \tau \\ 0 & \text{elsewhere.}^a \end{cases}$$

The actual time W during which an individual is under observation is the shorter of the two, X_1 and X_2:

$$W = \min(X_1, X_2) = \text{actual life.}$$

One computes its survival function

$$\begin{aligned} \bar{F}_W(w) &= P(W > w) \\ &= P(X_1 > w, Y_2 > w) \\ &= \int_w^\tau \int_w^\tau \left(\frac{2x_1}{\tau^2}\right)\left(\frac{1}{\tau}\right) dx_2\, dx_1 \\ &= \frac{(\tau - w)(\tau^2 - w^2)}{\tau^3} \qquad \text{for } 0 \leqslant w < \tau \end{aligned}$$

while one clearly has

$$\begin{aligned} \bar{F}_W(w) &= 1 \qquad \text{for } w < 0 \\ \bar{F}_W(w) &= 0 \qquad \text{for } w > \tau. \end{aligned}$$

Since, for practical purposes, only the actual duration of observation W and the "cause" of termination (condition C or dropping out) can be observed, we consider the two-dimensional random variable

$$(W, J)$$

where $J = 1$ when C occurs first, i.e., if $X_1 \leqslant X_2$, and $J = 2$ when the patient drops out before C occurs, i.e., if $X_2 < X_1$.

[a] These probability densities were chosen strictly for the simplicity of computations, without any pretense of corresponding to a realistic situation.

The joint probability distribution of (W, J) can be computed as follows: for $0 \leqslant w \leqslant \tau$ we have

$$P(W > w, J = 1) = P(w < X_1 \leqslant X_2)$$

$$= \int_w^\tau \int_{x_1}^\tau \left(\frac{2x_1}{\tau^2}\right)\left(\frac{1}{\tau}\right) dx_2\, dx_1$$

$$= \frac{1}{3} - \left(\frac{w}{\tau}\right)^2 + \frac{2}{3}\left(\frac{w}{\tau}\right)^3,$$

and similarly

$$P(W > w, J = 2) = P(w < X_2 < X_1)$$

$$= \int_w^\tau \int_{x_2}^\tau \left(\frac{1}{\tau}\right)\left(\frac{2x_1}{\tau^2}\right) dx_1\, dx_2$$

$$= \frac{2}{3} - \left(\frac{w}{\tau}\right) + \frac{1}{3}\left(\frac{w}{\tau}\right)^3,$$

while for $w < 0$ one has

$$P(W > w, J = 1) = \frac{1}{3}$$

$$P(W > w, J = 2) = \frac{2}{3}$$

and for $w > \tau$ one has

$$P(W > w, J = 1) = P(W > w, J = 2)$$

$$= 0.$$

The marginal probabilities of termination due to either of the two risks are, therefore:

$$P(\text{termination due to } C) = P(J = 1)$$

$$= P(W > 0, J = 1)$$

$$= \frac{1}{3},$$

$$P(\text{termination due to dropout}) = P(J = 2)$$

$$= P(W > 0, J = 2)$$

$$= \frac{2}{3}.$$

The conditional probability distribution of the time to termination, given that it occurs due to condition C, can be stated in form of the survival function

$$P(W > w|J = 1) = \frac{P(W > w, J = 1)}{P(J = 1)}$$

$$= 1 - 3\left(\frac{w}{\tau}\right)^2 + 2\left(\frac{w}{\tau}\right)^3 \qquad \text{for } 0 \leqslant w \leqslant \tau,$$

and a random variable Y_1 that has this probability distribution, i.e., has the survival function

$$\bar{F}_{Y_1}(y_1) = 1 - 3\left(\frac{y_1}{\tau}\right)^2 + 2\left(\frac{y_1}{\tau}\right)^3 \qquad \text{for } 0 \leqslant y_1 \leqslant \tau$$

is customarily referred to as the "crude" life due to the first risk.

Similarly, the conditional probability distribution of the time to termination, given it occurs due to dropping out, is described by

$$P(W > w|J = 2) = \frac{P(W > w, J = 2)}{P(J = 2)}$$

$$= 1 - \frac{3}{2}\left(\frac{w}{\tau}\right) + \frac{1}{2}\left(\frac{w}{\tau}\right)^3$$

and the random variable Y_2 that has this probability distribution, i.e., has the survival function

$$\bar{F}_{Y_2}(y_2) = P(Y_2 > y_2)$$

$$= 1 - \frac{3}{2}\left(\frac{y_2}{\tau}\right) + \frac{1}{2}\left(\frac{y_2}{\tau}\right)^3 \qquad \text{for } 0 \leqslant y_2 \leqslant \tau$$

is traditionally called the "crude" life due to the second risk.

2.3. GENERAL THEORY: NET LIVES

In general there may be any number k of "modes of failure" or "risks," which will be denoted by $R_1, R_2, \cdots, R_k$. We assume that an individual organism, or system, or device, fails at the earliest occurrence of one of these modes of failure. To each of the risks corresponds a "net" or "potential" life; we shall interpret these net lives as a k-dimensional random vector

$$(X_1, X_2, \cdots, X_k) = \underline{X}$$

the coordinate X_j being the time to the occurrence of R_j. It should be noted that these net lives $X_1, \cdots, X_k$ may be dependent random variables.

Since the system fails at the realization of the earliest risk, the time to failure for the system is

$$W = \min (X_1, X_2, \cdots, X_k), \tag{8}$$

called the "system-life" or "actual life."

We shall assume that the net lives have a joint probability density

$$f(\underline{x}) = f(x_1, x_2, \cdots, x_k),$$

so that the joint probability distribution function may be written as

$$\begin{aligned} F_{\underline{X}}(\underline{x}) &= F_{\underline{X}}(x_1, \cdots, x_k) \\ &= P(X_1 \leqslant x_1, \cdots, X_k \leqslant x_k) \\ &= \int_0^{x_1} \cdots \int_0^{x_k} f(s_1, \ldots, s_k)\, ds_k \ldots ds_1 \end{aligned}$$

and the joint survival function as

$$\begin{aligned} \overline{F}_{\underline{X}}(\underline{x}) &= \overline{F}_{\underline{X}}(x_1, \ldots, x_k) \\ &= P(X_1 > x_1, \ldots, X_k > x_k) \\ &= \int_{x_1}^{\infty} \cdots \int_{x_k}^{\infty} f(s_1, \ldots, s_k)\, ds_k \ldots ds_1. \end{aligned} \tag{9}$$

Clearly, the marginal probability distribution function of the net life X_i is

$$\begin{aligned} F_{X_i}(x_i) &= P(X_i \leqslant x_i) \\ &= F_{\underline{X}}(\infty, \cdots, \infty, x_i, \infty, \cdots, \infty) \end{aligned}$$

and the marginal survival function

$$\begin{aligned} \overline{F}_{X_i}(x_i) &= P(X_i > x_i) \\ &= \overline{F}_{\underline{X}}(0, \cdots, 0, x_i, 0, \cdots, 0), \end{aligned}$$

with x_i in either case at the ith place in the lower expression.

In view of equation (8), the survival function for the system life W is

$$\begin{aligned} \overline{F}_W(w) &= P(W > w) \\ &= \overline{F}_{\underline{X}}(w, w, \cdots, w). \end{aligned}$$

From now on it will be assumed that, whenever the system fails, one can observe only its actual life and the mode of failure, i.e., the value of the two-dimensional random variable

$$(W, J)$$

where W is the system life and $J = j$ when

$$W = \min (X_1, X_2, \cdots, X_k)$$

$$= X_j, \qquad j = 1, 2, \cdots, k.$$

Since, under our assumptions,

$$P(X_i = X_j) = 0 \qquad \text{for } i \neq j,$$

so that the probability of two or more net lives being equal is zero, the value of J is uniquely determined with probability one.

The joint probability distribution of (W, J) can be expressed as

$$F_{W,J}(w, j) = P(W \leqslant w, J = j)$$

$$= P(X_j \leqslant w, X_j < X_i \text{ for } i \neq j)$$

$$= \int_0^w \left[\int_{x_j}^{\infty} \cdots \int_{x_j}^{\infty} f(x_1, \cdots, x_k) \prod_{r \neq j} dx_r \right] dx_j \qquad \text{for } j = 1, \cdots, k. \tag{10}$$

It follows that the probability distribution function for the actual life W is

$$F_W(w) = P(W \leqslant w)$$

$$= \sum_{j=1}^{k} F_{W,J}(w, j).$$

and the corresponding survival function

$$\bar{F}_W(w) = 1 - \sum_{j=1}^{k} F_{W,J}(w, j).$$

One also has for the marginal probability that failure will occur at any time due to R_j the expression

$$P(J = j) = F_{W,J}(+\infty, j)$$

$$= \int_0^{\infty} \left[\int_{x_j}^{\infty} \cdots \int_{x_j}^{\infty} f(x_1, \cdots, x_k) \prod_{i \neq j} dx_i \right] dx_j. \tag{11}$$

2.4. GENERAL THEORY: CRUDE LIVES

Since we assume that only W and J, the actual life and the mode of failure, are observable, it is of interest to consider the conditional probability distribution of W given that $J = j$

$$P(W \leqslant w \mid J = j) = \frac{P(W \leqslant w, J = j)}{P(J = j)} \qquad \text{for } j = 1, 2, \cdots, k.$$

It has been customary to consider for each value of j a random variable Y_j which has this probability distribution and hence corresponds to the observable time to failure due to risk R_j. This random variable is usually called the "crude" life for risk R_j; its probability distribution function is

$$\begin{aligned} F_{Y_j}(y_j) &= P(Y_j \leqslant y_j) \\ &= \frac{P(W \leqslant y_j, J = j)}{P(J = j)} \\ &= \frac{F_{W,J}(y_j, j)}{P(J = j)} \end{aligned} \tag{12}$$

where the numerator can be computed according to equation (10) and the denominator according to equation (11).

We shall consider the probability

$$\begin{aligned} Q_{Y_j}(t) &= P(W > t, J = j) \\ &= P(X_j > t, \text{ and } X_j < X_i \text{ for } i \neq j) \\ &= P(J = j \text{ and } Y_j > t) \\ &= P(J = j) - F_{W,J}(t, j) \qquad \text{for } j = 1, \cdots. \end{aligned} \tag{13}$$

This is the probability of survival beyond age t and then failure due to R_j. It should be noted that

$$Q_{Y_j}(0) = P(J = j \text{ and } Y_j > 0) \leqslant P(J = j),$$

hence $Q_{Y_j}(0)$ need not be 1, so that in general $Q_{Y_j}(t)$ is not a survival function according to the definition in section 2.1.

2.5. AN IDENTITY

For $j = 1, 2, \cdots, k$ one has

$$\frac{d}{dt} Q_{Y_j}(t) = \frac{\partial}{\partial x_j} \bar{F}_{\underline{X}}(x_1, \ldots, x_k)\Big|_{x_1 = x_2 \ldots = x_k = t} \tag{14}$$

Proof: Without loss of generality, we prove this identity for $j = 1$. From equation (13)

$$Q_{Y_1}(t) = P(X_1 > t \text{ and } X_1 < X_i \quad \text{for } i \neq 1)$$

$$= \int_t^\infty \int_{x_1}^\infty \cdots \int_{x_1}^\infty f(x_1, s_2, \cdots, s_k)\, ds_k \cdots ds_2\, ds_1 \tag{15}$$

hence,

$$\frac{d}{dt}\, Q_{Y_1}(t) = -\int_t^\infty \cdots \int_t^\infty f(t, s_2, \cdots, s_k)\, ds_k \cdots ds_2 .$$

But from equation (9) one obtains for the right-hand side of equation (14)

$$\frac{\partial}{\partial x_1}\, \bar{F}_{\underline{X}}(x_1, \ldots, x_k)\Big| x_1 = x_2 \ldots = x_k = t$$

$$= -\int_{x_2}^\infty \cdots \int_{x_k}^\infty f(x_1, s_2, \cdots, s_k)\, ds_k \cdots ds_2 \mid x_1 = \cdots = x_k = t$$

$$= -\int_t^\infty \cdots \int_t^\infty f(t, s_2, \cdots, s_k)\, ds_k \cdots ds_2$$

which concludes the proof.

Identity (14) is due to Tsiatis (1975).[15] Under the assumption of independent net lives, Berman (1963)[16] gave an identity equivalent with equation (15) and explored its consequences, as described in section 2.6.2.

2.6. THE PROBLEM OF IDENTIFIABILITY

2.6.1. Formulation of Problem

Since only the actual life W and the mode of failure are observable, the question arises how much information can such observations yield about the probability distribution of net lives. One would like to learn as much as possible about that probability distribution since, as we have seen, all other probability distributions of our theory follow from it.

If there are sufficiently many observations of W and of J, then it should be possible to estimate very closely the probabilities $P(J = j)$ of the modes of failure and $F_{Y_j}(t)$ for the crude lives. Appropriate estimation techniques will be discussed in section 3; for the time being we shall *assume* that practically unlimited numbers of observations were available so *that* $P(J = j)$ *and* $F_{Y_j}(t)$ *for* $j = 1, 2, \cdots, k$ *are known.*

The problem of identifiability can now be stated: If $P(J = j)$ and $F_{Y_j}(t)$ are given for $j = 1, 2, \cdots, k$, does this determine the joint probability distribution of the net lives $X_1, X_2, \cdots, X_k$?

2.6.2. The Case of Independent Net Lives

We have observed that the net lives need not be independent. In fact, the assumption of their independence is in many practical situations obviously unrealistic.

In this section, however, we make the assumption that $X_1, X_2, \cdots, X_k$ are independent, that is

$$f_{\underline{X}}(\underline{x}) = \prod_{j=1}^{k} f_{X_j}(x_j)$$

or equivalently

$$\bar{F}_{\underline{X}}(\underline{x}) = \prod_{j=1}^{k} \bar{F}_{X_j}(x_j).$$

From this, using the definition (6) of the hazard function, one has

$$\frac{\partial}{\partial x_i} \bar{F}_{\underline{X}}(\underline{x}) = -f_{X_i}(x_i) \prod_{j \neq 1} \bar{F}_{X_j}(x_j).$$

$$= -\lambda_{X_i}(x_i)\bar{F}_{\underline{X}}(\underline{x}) \tag{16}$$

and identity (14) yields

$$\frac{d}{dt} Q_{Y_i}(t) = \frac{\partial}{\partial x_i} \bar{F}_{\underline{X}}(x_1, \cdots, x_k) \Big| x_1 = x_2 = \cdots = x_k = t$$

$$= -\lambda_{X_i}(t)\bar{F}_{\underline{X}}(t, \cdots, t). \tag{17}$$

Using equation (7) we have

$$\bar{F}_{X_i}(t) = \exp\left[-\int_0^t \lambda_{X_i}(s)\, ds\right]$$

and with the notation

$$\sum_{j=1}^{k} \lambda_{X_j}(x) = \lambda(x)$$

we may write

$$\overline{F}_{\underline{X}}(t, \cdots, t) = \prod_{j=1}^{k} \overline{F}_{X_j}(t)$$

$$= \exp\left[- \int_0^t \lambda(s)\, ds \right]. \tag{18}$$

Since the survival function for the actual life W is

$$\overline{F}_W(t) = P(X_1 > t, \cdots, X_k > t)$$

$$= \overline{F}_{\underline{X}}(t, \cdots, t),$$

we see that $\lambda(x) = \lambda_w(x)$, and we note this fact as follows:

Lemma.—If the net lives are independent, then the hazard rate of the actual life W is

$$\lambda_W(x) = \lambda(x) = \sum_{j=1}^{k} \lambda_{X_j}(x). \tag{19}$$

Together with equation (17) this yields

$$\frac{d}{dt}\, Q_{Y_i}(t) = - \lambda_{X_i}(t) \exp\left[- \int_0^t \lambda(s)\, ds \right] \qquad \text{for } i = 1, 2, \cdots, k. \tag{20}$$

Now, when the $P(J = j)$ and $F_{Y_j}(y_j)$ are known for $j = 1, \cdots, k$, then by equations (12) and (13) $Q_{Y_j}(t)$ are known, and relationship (20) is a system of k equations with k unknown functions $\lambda_{X_i}(t)$. To solve it, we sum equations (20) over i

$$\frac{d}{dt} \sum_{i=1}^{k} Q_{Y_i}(t) = - \lambda(t) \exp\left[- \int_0^t \lambda(s)\, ds \right]$$

hence,

$$\sum_{i=1}^{k} Q_{Y_i}(t) = \exp\left[- \int_0^t \lambda(s)\, ds \right] + c.$$

In view of

$$\sum_{i=1}^{k} Q_{Y_i}(0) = \sum_{i=1}^{k} P(I = i) = 1$$

we have $c = 0$, and from equation (20)

$$\lambda_{X_i}(t) = -\frac{\left(\frac{d}{dt}\right) Q_{Y_i}(t)}{\sum_{i=1}^{k} Q_{Y_i}(t)} \qquad \text{for } i = 1, \cdots, k, \tag{21}$$

so that, knowing the $Q_{Y_i}(t)$ for $i = 1, \cdots, k$, we can compute all hazard rates $\lambda_{X_i}(t)$, and by equation (7) all $\bar{F}_{X_i}(t)$ and $F_{X_i}(t)$.

To summarize, if the net lives are independent, then the probabilities $P(W \leqslant w, J = j)$, or the probabilities $P(J = j)$ and $F_{Y_j}(t)$ for $j = 1, \cdots, k$, determine all $\bar{F}_{X_j}(t)$ and all $F_{X_j}(t)$, hence, also the joint probability distribution of all net lives $X_1, \cdots, X_k$. This result is due to Berman.[16]

2.6.3. The Case of Possibly Dependent Net Lives

The assumption of independence of the net lives was used to obtain equations (16) and (17), and to derive formulas (21) which show how, given the probabilities $Q_{Y_j}(t)$ for the crude lives (or, equivalently, the probabilities $P(J)$ and $F_{Y_j}(Y_j)$), one can compute the probability distributions of the net lives.

This suggests two questions:

(a) Can formula (21) be used to obtain the distributions of net lives from those of crude lives if it is not known that the net lives are independent?
(b) If it is not assumed that $X_1, \cdots, X_k$ are independent, do the $P(J)$ and $F_{Y_j}(y_j)$, $j = 1, \cdots, k$ determine the joint distribution of $X_1, \cdots, X_k$?

The following example, given by Tsiatis,[15] shows that the answer to question (a) is negative.

Example C.—Let the joint survival function of (X_1, X_2) be of the form

$$\bar{F}_{\underline{X}}(x_1, x_2) = \exp(-\lambda x_1 - \mu x_2 - \gamma x_1 x_2) \tag{22}$$

where $\lambda > 0, \mu > 0, 0 \leqslant \gamma \leqslant \lambda\mu$. Clearly, X_1, X_2 are independent if and only if $\gamma = 0$. The marginal survival functions for the net lives are

$$\begin{aligned} \bar{F}_{X_1}(x_1) &= \exp(-\lambda x_1) \\ \bar{F}_{X_2}(x_2) &= \exp(-\mu x_2) \end{aligned} \tag{23}$$

and do not depend on γ.

By equation (14)

$$\frac{d}{dt} Q_{Y_1}(t) = -(\lambda + \gamma t) \exp(-\lambda t - \mu t - \gamma t^2)$$

$$\frac{d}{dt} Q_{Y_2}(t) = -(\mu + \gamma t) \exp(-\lambda t - \mu t - \gamma t^2).$$

Integrating each of these identities one obtains for $Q_{Y_1}(t)$ and $Q_{Y_2}(t)$ expressions that depend on all three parameters λ, μ, γ in such a manner that, substituted in the right-hand side of equation (21),

they would lead to hazard rates $\lambda_{X_1}(t)$, $\lambda_{X_2}(t)$ which depend on all three parameters and hence on γ. These hazard rates would, therefore, differ from the correct ones which correspond to equations (23) and which are independent of γ. Thus, it is shown that equation (21) cannot be used in the case of dependent net lives.

The next example, due to Rose,[17] answers question (b), also in the negative.

Example D.—Let $k = 2$ and

$$f_{X_1,X_2}(x_1, x_2) = \begin{cases} 1 & \text{for } 0 \leqslant x_1 \leqslant 1 \text{ and } 0 \leqslant x_2 \leqslant 1 \\ 0 & \text{elsewhere,} \end{cases}$$

i.e., let the net lives X_1, X_2 be independent and each uniformly distributed on $[0, 1]$.

Consider now the family of two-dimensional probability densities, obtained by modifying $f_{X_1,X_2}(x_1, x_2)$ in a square with sides of length a, with $0 \leqslant a < (1/2)$ (see figure 3):

$$ {}_af_{X_1,X_2}(x_1, x_2) = \begin{cases} 0 & \text{when } 0 \leqslant x_1 < \frac{a}{2} \text{ and } 1 - \frac{a}{2} < x_2 \leqslant 1 \\ 0 & \text{when } \frac{a}{2} \leqslant x_1 < a \text{ and } 1 - a \leqslant x_2 < 1 - \frac{a}{2} \\ 2 & \text{when } 0 \leqslant x_1 < \frac{a}{2} \text{ and } 1 - a \leqslant x_2 < 1 - \frac{a}{2} \\ 2 & \text{when } \frac{a}{2} \leqslant x_1 < a \text{ and } 1 - \frac{a}{2} \leqslant x_2 < 1 \\ f_{X_1,X_2}(x_1, x_2) & \text{everywhere else.} \end{cases} \tag{24}$$

Clearly $f_{X_1,X_2}(x_1, x_2)$ is contained in this family, since

$$ {}_0f_{X_1,X_2}(x_1, x_2) = f_{X_1,X_2}(x_1, x_2). $$

One verifies that, for every ${}_af_{X_1,X_2}(x_1, x_2)$, the net lives have marginal probability densities which are uniform on the unit interval

$$ {}_af_{X_j}(x_j) = \begin{cases} 1 & \text{for } 0 \leqslant x_j \leqslant 1 \\ 0 & \text{elsewhere} \end{cases} \qquad \text{for } j = 1, 2, $$

hence are the same for all a, $0 \leqslant a < 1/2$. For $a > 0$ these net lives are not independent, for in the entire square $0 \leqslant x_1 < a/2$, $1 - a/2 < x_2 \leqslant 1$ one has

$$ {}_af_{X_1,X_2}(x_1, x_2) = 0 \neq {}_af_{X_1}(x_1) \cdot {}_af_{X_2}(x_2) = 1 $$

when $0 < a < 1/2$.

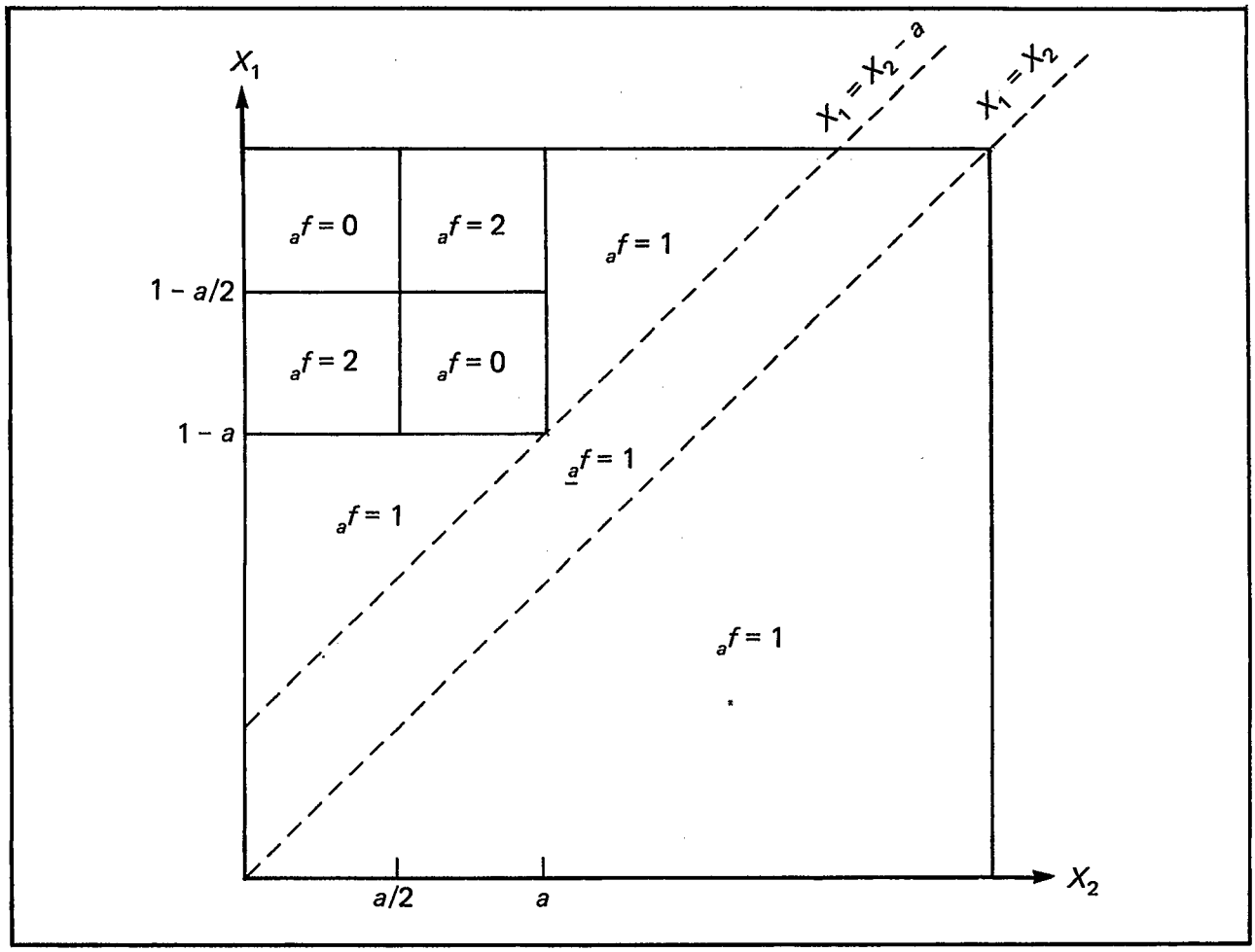


Figure 3. Probability densities used in example D.

The joint survival function of the net lives is

$$ {}_a\overline{F}_{X_1,X_2}(x_1, x_2) = (1 - x_1)(1 - x_2) \qquad \text{for } 0 \leqslant x_2 \leqslant x_1 + a \leqslant 1, \tag{25} $$

and is equal to other expressions in other parts of the unit square. However, to use equation (14) for computing

$$ \frac{d}{dt}\, Q_{Y_1}(t) \text{ and } \frac{d}{dt}\, Q_{Y_2}(t), $$

it is enough to know the joint survival function in an open set containing the diagonal $x_1 = x_2$ and equation (25) provides this information. Hence, by equations (14) and (25), we have

$$ \frac{d}{dt}\, Q_{Y_1}(t) = -(1 - t) $$

and

$$Q_{Y_1}(t) = \frac{t^2}{2} - t + Q_{Y_1}(0)$$

$$= \frac{t^2}{2} - t + P(J = 1) \qquad \text{for } 0 \leqslant t \leqslant 1.$$

Since $P(J = 1) = P(X_1 < X_2) = 1/2$ independently of a, we have by equations (12) and (13) for the crude life Y_1

$$F_{Y_1}(y_1) = 2y_1 - y_1^2 \qquad \text{for } 0 \leqslant y_1 \leqslant 1. \tag{26}$$

By an analogous argument one obtains

$$F_{Y_2}(y_2) = 2y_2 - y_2^2 \qquad \text{for } 0 \leqslant y_2 \leqslant 1. \tag{27}$$

Obviously

$$F_{Y_j}(y_j) = \begin{cases} 0 & \text{for } y_j < 0 \\ 1 & \text{for } y_j > 1, j = 1, 2. \end{cases}$$

In example D, $F_{Y_1}(y_1)$ and $F_{Y_2}(y_2)$ as well as $P(J = 1)$, $P(J = 2)$ are the same for all values $0 \leqslant a < 1/2$, and the same is, therefore, true of $F_{W,J}(w, j)$. The knowledge of these probabilities, therefore, does not identify the joint probability distribution of the net lives X_1, X_2, and the answer to question (b) is negative.

2.6.4. An Interpretation of the Identifiability Problem: Some Inequalities

We limit ourselves to the case of two competing risks, $k = 2$, and consider the probabilities defined in equations (9) and (13) which, in our case, we denote simply by

$$\bar{F}(x_1, x_2) = P(X_1 > x_1, X_2 > x_2) \tag{28}$$

and

$$\begin{aligned} Q_1(x_1) &= P(X_1 > x_1 \text{ and } X_1 < X_2) \\ Q_2(x_2) &= P(X_2 > x_2 \text{ and } X_2 < X_1). \end{aligned} \tag{29}$$

We assume as before that $Q_1(x_1)$ and $Q_2(X_2)$ can be estimated, hence will be considered as known, and restate the problem of identifiability in the form: To what extent do the functions $Q_1(x_1)$ and $Q_2(x_2)$ determine the joint survival function $\bar{F}(x_1, x_2)$ of the net lives, or at least one of the marginal survival functions $P(X_1 > x_1) = \bar{F}_{X_1}(x_1) = \bar{F}(x_1, 0)$, $P(X_2 > x_2) = \bar{F}_{X_2}(x_2) = \bar{F}(0, x_2)$?

As indicated on figure 4, $\bar{F}(x_1, x_2)$ is the integral of the joint probability density $f(x_1, x_2)$ over the region contained in the right angle between bold lines:

$$\bar{F}(x_1, x_2) = \int\int f(x_1, x_2)\, dx_1\, dx_2,$$

while $Q_1(x_1)$ is the integral over the interior of the 45° angle bounded by broken lines

$$Q_1(x_1) = \int\int f(x_1, x_2)\, dx_1\, dx_2$$

and $Q_2(x_2)$ is the integral over the interior of the 45° angle bounded by dotted lines

$$Q_2(x_2) = \int\int f(x_1, x_2)\, dx_1\, dx_2.$$

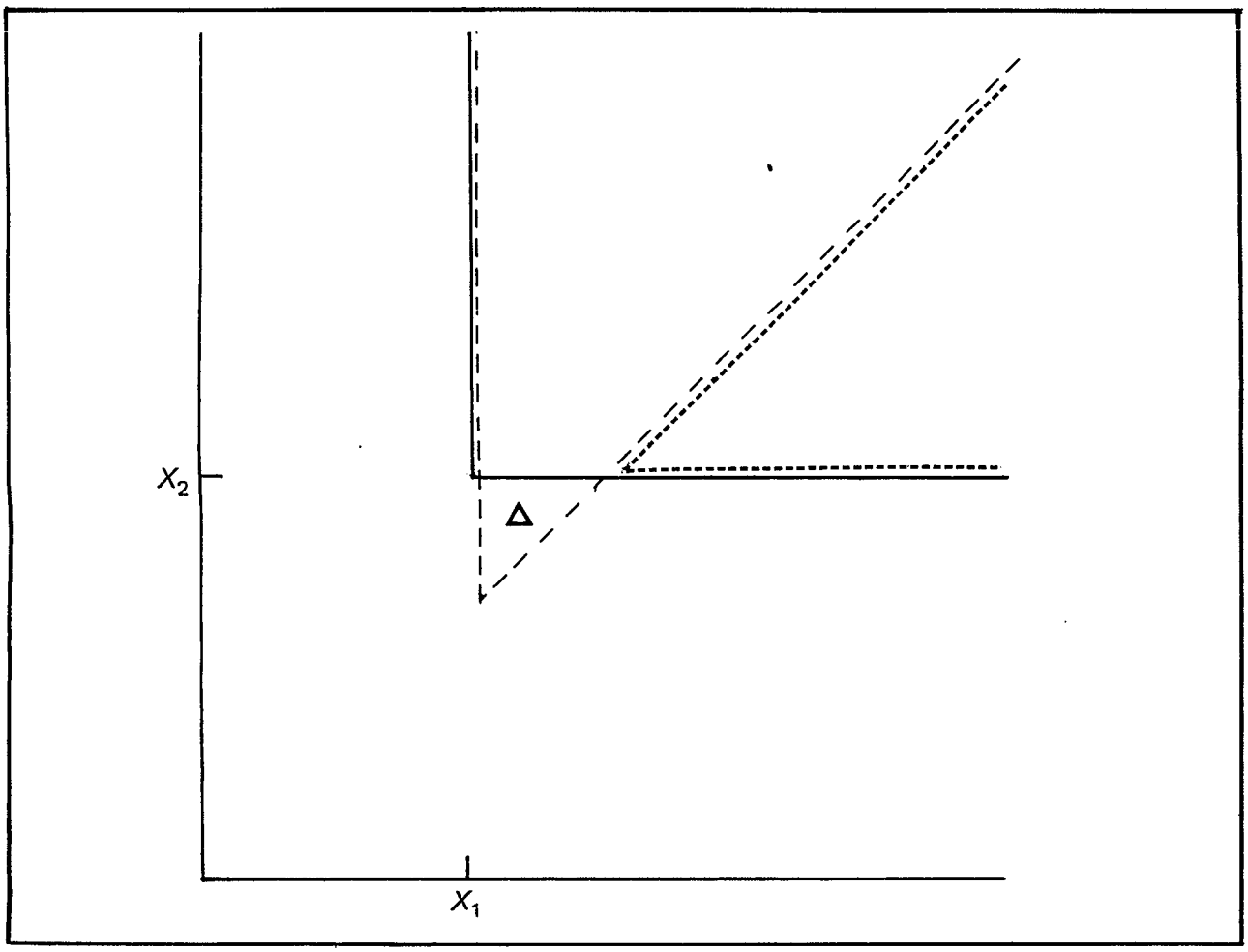


Figure 4. Regions of integration considered in equations (30) and (31).

By inspection one has

$$\bar{F}(x_1, x_2) = Q_1(x_1) + Q_2(x_2) - P(\Delta) \tag{30}$$

where $P(\Delta)$ is the probability of (X_1, X_2) falling into the triangle denoted by Δ in figure 4 or, more explicitly,

$$\bar{F}(x_1, x_2) = Q_1(x_1) + Q_2(x_2) - P(x_1 < X_1 < X_2 < x_2) \tag{31}$$

when $x_1 < x_2$ (as in figure 4) or

$$\bar{F}(x_1, x_2) = Q_1(x_1) + Q_2(x_2) - P(x_2 < X_2 < X_1 < x_1) \tag{32}$$

when $x_2 < x_1$. As an immediate consequence one obtains the inequalities

$$Q_1[\max(x_1, x_2)] + Q_2[\max(x_1, x_2)] \leqslant \bar{F}(x_1, x_2) \leqslant Q_2(x_2) + Q_1(x_1). \tag{33}$$

Writing

$Q_1(0) = P(X_1 < X_2)$ = probability that failure occurs due to the first risk = p_1

$Q_2(0) = P(X_2 < X_1)$ = probability that failure occurs due to the second risk = p_2

and recalling that survival probabilities for the net lives are

$$\bar{F}_{X_1}(x_1) = \bar{F}(x_1, 0),$$

$$\bar{F}_{X_2}(x_2) = \bar{F}(0, x_2),$$

one obtains from equation (33) the further inequalities

$$Q_1(x_1) + Q_2(x_1) \leqslant \bar{F}_{X_1}(x_1) \leqslant Q_1(x_1) + p_2 \tag{34}$$

$$Q_1(x_2) + Q_2(x_2) \leqslant \bar{F}_{X_2}(x_2) \leqslant p_1 + Q_2(x_2). \tag{35}$$

The inequalities (32), (33), and (34) were obtained in 1975 by Arthur V. Peterson, Jr.,[18] by a formal argument.

The identities (30), (31), and (32) shed additional light on the question of identifiability: If only crude lives are observable, the probabilities $P(\Delta)$ cannot be estimated, and this is the reason why $Q_1(x_1)$ and $Q_2(x_2)$ do not determine $\bar{F}(x_1, x_2)$ in the general case. It may be worth noting, however, that according to identity (30)

$$\bar{F}(x, x) = Q_1(x) + Q_2(x) \tag{36}$$

i.e., the knowledge of both crude-life distributions determines the survival function $\bar{F}(x, x) = P(W > x)$ of W, the actual life, without the assumption of independence.

2.7. SUMMARY AND COMMENTS ON IDENTIFIABILITY

We have seen that the joint probability distribution of the net lives $X_1, \ldots, X_k$, as given by the survival function $\bar{F}_X(x_1, \ldots, x_k)$, determines the probabilities of the crude lives, i.e., all

$$P(J = j) \text{ and } F_{Y_j}(y_j) \qquad \text{for } j = 1, \ldots, k$$

or, equivalently, all

$$P(W \leqslant w, J = j) = F_{W,J}(w,j) \qquad \text{for } j = 1, \ldots, k.$$

The converse is true if the net lives $X_1, \ldots, X_k$ are independent. In general however, i.e., when it is not assumed that the net lives are independent, the probability distributions of the crude lives do not determine the probability distribution of the net lives.

This brings us to a serious practical problem. In competing risk situations, the only observable quantities are the actual life W and the mode of failure J (i.e., the crude lives), and the best one can hope for is to obtain good estimates of their probability distributions. How does one from then on draw conclusions about the net lives in cases where the assumption of independence is unrealistic? Clearly, the knowledge of the probability distribution of W and J will have to be supplemented by some further information, either stated in form of further assumptions, or drawn from additional observations.

To see what some additional assumptions can contribute, let us return to example C, section 2.6.3. In that example we assumed that $Q_{Y_1}(t)$ and $Q_{Y_2}(t)$ were known and we attempted, unsuccessfully, to determine $\bar{F}_{X_1, X_2}(x_1, x_2)$ by applying equation (21). We did not, however, make use of any assumptions on the form of $\bar{F}_{X_1, X_2}(x_1, x_2)$. Let us now assume again that $Q_{Y_1}(t)$, $Q_{Y_2}(t)$ are known and use the *additional knowledge* that $\bar{F}_{X_1, X_2}(x_1, x_2)$ is of the form given by equation (22). Now all one needs to determine $\bar{F}_{X_1, X_2}(x_1, x_2)$ are the values of the three parameters λ, μ, and γ. But this should be quite feasible: If $Q_{Y_1}(t)$ is a known function of t, then the left side in the identity

$$\frac{d}{dt} Q_{Y_1}(t) = -(\lambda + \gamma t) \exp(-\lambda t - \mu t - \gamma t^2) \tag{37}$$

is also a known function. Entering any three different values t_1, t_2, t_3 in equation (37), one obtains three equations that then can be solved for λ, μ, and γ. In fact, knowing the probability distribution for only *one crude life,* together with the assumption that $\bar{F}_{X_1, X_2}(x_1, x_2)$ belongs to the parametric family of equation (22), is sufficient to determine completely the joint distribution of the net lives.

This example points to *one effective way for overcoming the difficulty* due to the fact that, in general, the probability distribution of net lives is not determined by the probability distributions of crude lives: *One assumes that* $\bar{F}_{\underline{X}}(\underline{x})$ *belongs to a specified parametric family.* This does, of course, raise another question: How does one choose a parametric model which is appropriate for the phenomena under investigation? In some cases a thorough knowledge of the physiological or technological processes underlying the functioning and failure of the systems may lead to the proper choice of such a parametric model. If there are no such guidelines, one may be tempted to choose a parametric model for its mathematical simplicity and manageability—a situation, and responsibility, not unfamiliar to those who deal with applied mathematics.[b]

---

[b] Professor A. J. Quinzi brought to our attention the 1977 and 1978 references Langberg, Proschan, and Quinzi,[19-21] in which the authors developed a general theory of converting models with dependent random variables

2.8. MULTIPLE DECREMENT TABLES VERSUS QUESTIONS ABOUT NET LIVES

When data on crude lives are available, it is important for some applications to investigate mainly the joint distributions of these crude lives. An important example of studies of this kind is the analysis of mortality data by different causes of death. Most of the work in this direction was done by actuaries who have developed techniques for constructing and using "multiple decrement tables."

This report will, from now on, deal only with questions that go in the direction initiated by Daniel Bernoulli and are of the following general kind: When empirical data are available on crude lives in a competing risks situation, what inferences can be drawn from these data about the net lives and their probability distributions, and what assumptions are needed to make such inferences possible?

3. ESTIMATION TECHNIQUES

3.1. CASE OF INDEPENDENT NET LIVES: THE KAPLAN-MEIER NONPARAMETRIC TECHNIQUE

3.1.1. Statement of Problem

We have seen that the probability distributions of the crude lives determine the probability laws of the net lives when one makes the assumption that the net lives are independent, or assumes a specified parametric model for the probability distribution of the net lives. We shall now discuss the case of independent net lives (hence will not assume any specific parametric form for their distribution) and present some of the known estimation procedures.

3.1.2. Definitions and Notations

In 1958, Kaplan and Meier[22] dealt with the following situation: Consider two competing risks which, to aid one's intuition, will be referred to as "death" and "loss from observation." For each individual there is a two-dimensional random variable (T, L) with

T = "time to death" = "life length"

L = "time to loss from observation" = "limit of observation,"

and T and L are assumed independent. At time $t = 0$ a cohort of $n(0)$ individuals is placed in observation. Each individual remains under observation until the earlier of the two events (death or loss), occurs so that observation records show only values of the actual life $W = \min(T, L)$ and the event that occurred at time W. Using these recorded observations, one wishes to estimate either one of the two net survival functions, say that of T

$$\bar{F}(t) = P(T > t). \tag{38}$$

---

into models with independent variables. As one of the applications, they obtained a set of assumptions which are different from assuming either independence or a parametric model, and are sufficient for a positive answer to the problem of identifiability.

The recorded observations determine the nonnegative, decreasing, integer-valued function defined by

$$n(t) = \text{number of individuals alive and in observation at time } t, \tag{39}$$

with the convention that a death occurring at t is already subtracted in computing $n(t)$, while a loss at t is not yet subtracted. In other words, deaths at t are treated as if they occurred slightly before t, and losses from observation at t are counted as if they occurred slightly after t. Some consequences of this convention are these properties of $n(t)$:

$$\begin{aligned} &n(t) \text{ is continuous at } t \text{ if neither death nor loss occurs at } t, \\ &n(t-0) - n(t) = \text{number of deaths occurring at } t, \\ &n(t) - n(t+0) = \text{number of losses occurring at } t. \end{aligned} \tag{40}$$

3.1.3. A General Estimation Procedure

A number of estimation techniques for $\bar{F}(t)$ can be obtained as special cases of the following general scheme.

Step (a).—The time axis is divided in some disjoint intervals

$$(0, u_1], (u_1, u_2], \ldots, (u_{j-1}, u_j], \ldots.$$

We introduce the notations

$$P_j = P(T > u_j) = \bar{F}(u_j), \tag{41}$$

and

$$p_j = \frac{P_j}{P_{j-1}} \ . \tag{42}$$

Step (b).—We decide to use some estimates $\tilde{p}_j$ for $p_j, j = 1, 2, \ldots$.

Step (c).—For any $t > 0$, we use as an estimate for $\bar{F}(t)$ the statistic

$$\tilde{H}(t) = \prod_{u_j < t} \tilde{p}_j. \tag{43}$$

In this general scheme, it is left open how the time intervals $(u_{j-1}, u_j]$ are chosen and what the estimates $\tilde{p}_j$ are. By choosing the division points u_j equidistant and assuming a specific parametric model for $\bar{F}(t)$, one can use standard curvefitting techniques to estimate the parameters, and the resulting $\tilde{p}_j$ and $\tilde{H}(t)$ will be actuarial estimates corresponding to the parametric model.

3.1.4. The Kaplan-Meier Estimates

The following specific procedures will now be used in carrying out steps (a) and (b) of the preceding section.

The recorded times of deaths and of losses are arranged in one increasing sequence of observed actual lives, $u_1 < u_2 < \cdots < u_j < \cdots$, and these u_j will be used as the endpoints of the intervals

$(u_{j-1}, u_j]$. For the sake of simplicity it will be assumed that only one of the events, death or loss, occurs at each u_j.

We write $n_j = n(u_j + 0)$ = number of individuals at risk (alive and in observation) immediately after u_j, and use as an estimate of the conditional probability p_j the statistic defined by

$$\hat{p}_0 = 1.$$

$$\hat{p}_j = \begin{cases} 1 & \text{if a loss occurred at } u_j, j = 1, 2, \ldots \\ \dfrac{n_{j-1} - 1}{n_{j-1}} & \text{if a death occurred at } u_j, j = 1, 2, \ldots. \end{cases} \tag{44}$$

The estimate (43) for $\bar{F}(t)$ now becomes

$$\hat{H}(t) = \prod_{u_j < t} \hat{p}_j. \tag{45}$$

This estimate will be referred to as the "Kaplan-Meier estimate for $\bar{F}(t)$." It clearly is a step function such that $\hat{H}(0) = 1$ and that $\hat{H}(t)$ remains unchanged when t increases across a value u_j at which a loss occurred, and $\hat{H}(t)$ decreases by a factor $(n_{j-1} - 1)/n_{j-1}$ when t crosses a value u_j at which a death occurred.

A small-size example of this procedure, obtained by a slight modification of an example given in 1958 by Kaplan and Meier (p. 464[22]) is presented in table A and figure 5.

3.1.5. Some Properties of the Kaplan-Meier Estimates

As the example of table 1 shows, the Kaplan-Meier estimate $\hat{H}(t)$ remains undetermined beyond the last actual life when the last event recorded is a loss. In our example we can only conclude that $0 \leqslant \hat{H}(t) \leqslant 21/80$ for $t > 12.1$. Had the last recorded event been a death, then we would have had $\hat{H}(t) = 0$ from then on.

When the data are large, some grouping may be desirable before carrying out the calculations leading to $\hat{H}(t)$. Several ways of doing this were suggested by Kaplan and Meier.

It may be important for some practical applications to note that the K-M procedure permits one to consider individuals entering into observation after the beginning of their lives. This is done by counting each such entrance as a negative loss, i.e., one records the time of entrance, increases the number of individuals in observation by 1, and otherwise treats the event as a loss. Such entering

Table A. Example of a Kaplan-Meier estimate

j	u_j	Events[1]	n_j	$\hat{p}_j$	$\hat{H}(u_j)$
0	0.0	---	8	1	1
1	0.8	Δ	7	7/8	7/8
2	1.0	λ	6	1	7/8
3	2.7	λ	5	1	7/8
4	3.1	Δ	4	4/5	7/10
5	5.4	Δ	3	3/4	21/40
6	7.0	λ	2	1	21/40
7	9.2	Δ	1	1/2	21/80
8	12.1	λ	0	1	21/80

[1]A death is indicated by Δ, a loss from observation by λ.

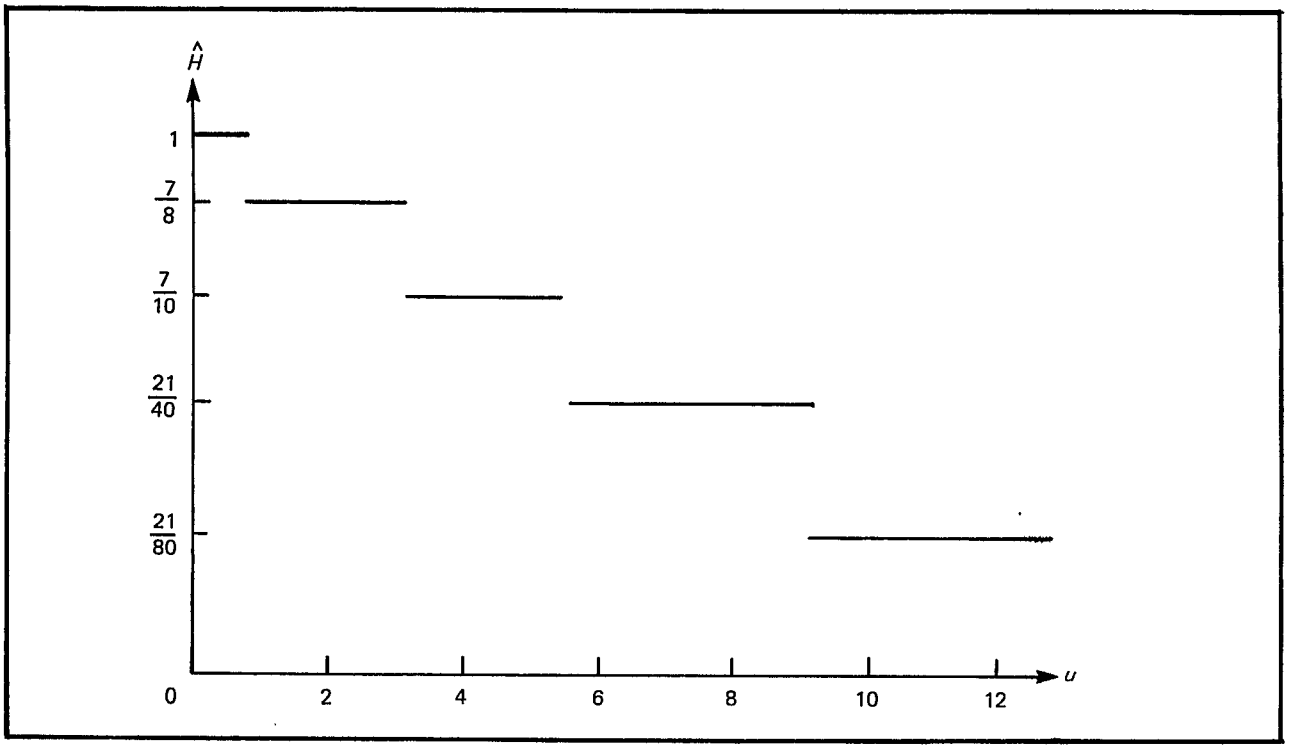


Figure 5. Graph of Kaplan-Meier estimate obtained in table A.

individuals can later on disappear from observation and will then be counted as ordinary losses. It should be stressed, however, that the assumption of independence of life length and time to loss is being used whether the loss is positive (individual drops out from observation) or negative (individual enters into observation).

Under the assumption of independence of T and L, Kaplan and Meier showed that $\hat{H}(t)$ is that member of the class of all survival functions which maximizes the likelihood function for the recorded values and that it is consistent. They also derived several approximate expressions for the variance of $\hat{H}(t)$. For these derivations, the reader is referred to the original paper. Much further theoretical research has been done on the asymptotic properties of the K-M estimates (Breslow and Crowley[23] in 1974, Crowley[24] in 1975, and Aalen[25] in 1976). A generalization to the case of more than two competing risks was introduced in 1972 by Hoel.[26]

3.2. SOME ACTUARIAL ESTIMATES AND THEIR RELATIONSHIPS TO KAPLAN-MEIER ESTIMATES

3.2.1. A Generalized K-M Procedure and Some Actuarial Estimates

The procedure described in section 3.1.4 can be obtained as a special case of the following more general way of carrying out the steps listed in section 3.1.3.

Step (a).—Given a recorded sequence of deaths and losses and the corresponding actual lives, we choose $u_1 < u_2 < \cdots < u_j < \cdots$ arbitrarily, but so that each interval $(u_{j-1}, u_j]$ contains only deaths, or only losses, not both. The u_j's need not be actual lives.

Step (b).—We introduce as estimates of the p_j the statistics defined by

$$\check{p}_0 = 1, \tag{46}$$

and, for $j = 1, 2, \ldots$, by

$$\check{p}_j = \begin{cases} \dfrac{n_{j-1} - \delta}{n_{j-1}} & \text{when } (u_{j-1}, u_j] \text{ contains } \delta \text{ deaths} \\ 1 & \text{when } (u_{j-1}, u_j] \text{ contains losses.} \end{cases} \tag{47}$$

Step (c).—We define

$$\check{H}(t) = \prod_{u_j < t} \check{p}_j. \tag{48}$$

This procedure clearly reduces to the Kaplan-Meier estimate when one chooses

$$u_1 < u_2 < \cdots u_j < \cdots$$

to be all actual lives.

For any interval $(a, b]$, $0 < a < b$, one has according to definition (48) the statistic

$$\check{p}_{a,b} = \prod_{a \leqslant u_j < b} \check{p}_j = \frac{\check{H}(b)}{\check{H}(a)} \tag{49}$$

as an estimate of the conditional probability

$$p_{a,b} = \frac{\bar{F}(b)}{\bar{F}(a)} .$$

To describe a class of approximations to $\check{p}_{a,b}$ which are often referred to as "actuarial" estimates, we consider an interval $(a, b]$ and the observed values:

$n = n(a + 0)$ = number of individuals under risk immediately after a

δ = number of deaths in $(a, b]$

λ = number of losses in $(a, b]$.

We assume that these three numbers are known, but the order in which deaths and losses follow each other is not known.

If all δ deaths happened to precede all λ losses then the estimate (49) could be obtained by choosing one division point and this estimate would be

$$\underline{p}_{a,b} = \frac{n - \delta}{n} . \tag{50}$$

Similarly, if all λ losses preceded all δ deaths, this estimate would have the value

$$\overline{p}_{a,b} = \frac{n - \lambda - \delta}{n - \lambda} \quad . \tag{51}$$

One readily verifies the inequality

$$\overline{p}_{a,b} \leqslant \underline{p}_{a,b} . \tag{52}$$

If the order of occurrence of deaths and losses were completely known, then $\check{p}_{a,b}$ could be computed according to (49).

None of the three quantities $\check{p}_{a,b}$, $\underline{p}_{a,b}$, $\overline{p}_{a,b}$ can be justified as an estimate of $p_{a,b}$ when only n, δ, λ are given but nothing is known about the order in which the events occurred, as is the case in many practical situations. In these cases, it has been customary to use as an approximation to $p_{a,b}$ the estimate

$$p^*_{a,b} = \frac{n - \lambda/2 - \delta}{n - \lambda/2} \tag{53}$$

sometimes referred to as the "adjusted-observed" acturial estimate[c] which satisfies the inequalities

$$\overline{p}_{a,b} \leqslant p^*_{a,b} \leqslant \underline{p}_{a,b} . \tag{54}$$

3.2.2. Conditions for $p^*_{a,b}$ Being Consistent

The estimate (53) is one of the most frequently used so-called "actuarial" estimates, and it is of some interest to determine if, as the number of observations increases, it tends in probability to $p_{a,b}$, i.e., if it is a consistent estimate of $p_{a,b}$. This matter was clarified in 1974 by Breslow and Crowley,[23] and the following is a presentation of their findings.

As before, we assume that the life T of an individual and the time L to loss from observation are independent random variables. Let their survival probabilities be

$$\overline{F}_T(t) = P(T > t) = 1 - F_T(t)$$

$$\overline{F}_L(t) = P(L > t) = 1 - F_L(t).$$

For any given interval $(a, b]$, we wish to estimate the conditional probability

$$p_{a,b} = \frac{\overline{F}_T(b)}{\overline{F}_T(a)}$$

by using estimate (53).

---

[c]The reader may compare this with Bernoulli's estimates for column (5) in his Table I (figure 2) which we mentioned in section 1.1.1, and Spurgeon's recommendation of formula (5).

Instead of intervals $(a, b]$ we may, without loss of generality, consider the intervals

$$I_\xi = (0,\xi], \xi > 0$$

with arbitrary ξ. We now define the following events (see figure 6):

$$\begin{aligned}
\delta_1 &= \{0 < T \leqslant \xi, L \geqslant \xi\} && \text{(individual dies in } I_\xi\text{, not due to be lost in } I_\xi\text{)},\\
\delta_2 &= \{0 < T \leqslant L < \xi\} && \text{(individual dies in } I_\xi\text{, before it is due for loss which is in } I_\xi\text{)},\\
\mu &= \{0 < L < \xi, L < T\} && \text{(individual is lost in } I_\xi \text{ and before its death)},\\
\gamma_1 &= \{0 < T, L \geqslant \xi\} && \text{(individual is not due for loss in } I_\xi\text{)},\\
\gamma_2 &= \{0 < T, L < \xi\} && \text{(individual is due for loss in } I_\xi\text{)}.
\end{aligned} \tag{55}$$

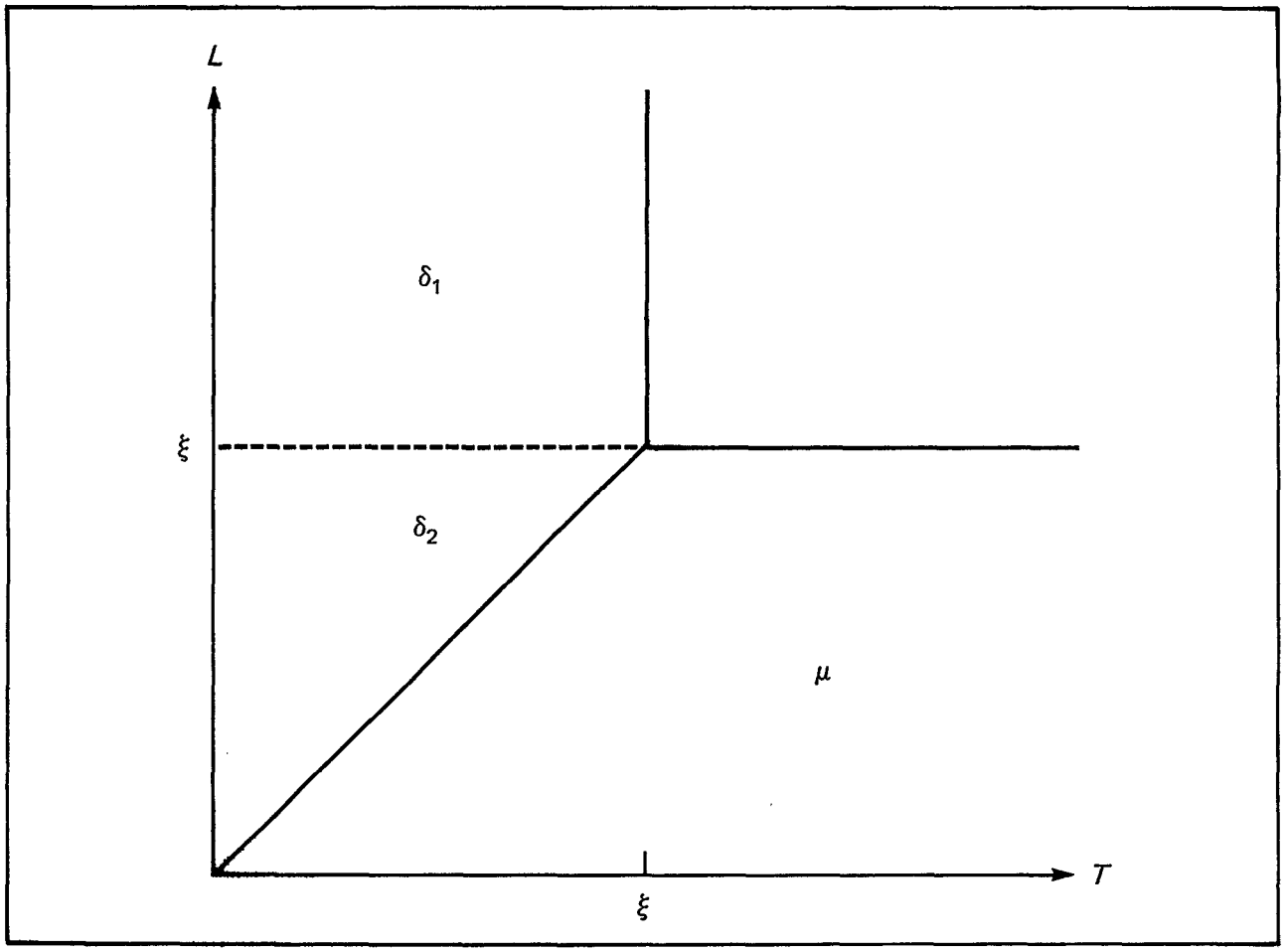


Figure 6. Diagram of events defined by formulas in (55).

Not all of these events are observable; however, the event μ is observable and so are the following unions

$\delta = \delta_1 \cup \delta_2$ (individual dies in I_ξ before getting lost)

$\gamma = \gamma_1 \cup \gamma_2$ (individual is alive and present, hence under risk, at 0).

Consider now the random variables.

$$\begin{aligned} N_1 &= \text{number of individuals for which } \gamma_1 \text{ occurs} \\ N_2 &= \text{number of individuals for which } \gamma_2 \text{ occurs} \\ D_1 &= \text{number of individuals for which } \delta_1 \text{ occurs} \\ D_2 &= \text{number of individuals for which } \delta_2 \text{ occurs} \\ M &= \text{number of individuals for which } \mu \text{ occurs.} \end{aligned} \tag{56}$$

One clearly has

$N = N_1 + N_2$ = number of individuals for which γ occurs, i.e., number under risk at 0

$D = D_1 + D_2$ = number of individuals for which δ occurs, i.e., whose deaths occur in I_ξ, before they are lost.

The random variables M, N, and D are observable. The probabilities of the events in (55) are

$$\begin{aligned} P(\delta_1) &= F_T(\xi)[1 - F_L(\xi)] \\ P(\delta_2) &= \int_0^\xi F_T(s)\, dF_L(s) \\ P(\mu) &= \int_0^\xi [1 - F_T(s)]\, dF_L(s) \\ P(\gamma_1) &= 1 - F_L(\xi) \\ P(\gamma_2) &= F_L(\xi) \end{aligned} \tag{57}$$

and, assuming the number N known, each of the random variables in (56) has binomial probability distribution with parameters N and the corresponding probability in (57).

The statistic (53) can now be written

$$p^*_{0,\xi} = \frac{N - M/2 - D}{N - M/2} = 1 - q^*_{0,\xi}$$

where

$$q^*_{0,\xi} = \frac{D}{N - M/2} = \frac{N_1}{N - M/2} \cdot \frac{D_1}{N_1} + \frac{N_2 - M/2}{N - M/2} \cdot \frac{D_2}{N_2 - M/2} \quad . \tag{58}$$

As $N \to \infty$, each of the quantities N_1/N, D_1/N, D_2/N, M/N tends in probability to the corresponding probability in (57) hence the right side of equation (58) tends to

$$\frac{1 - F_L(\xi)}{1 - (1/2)\int_0^\xi (1 - F_T)\, dF_L} \times \frac{F_T(\xi)[1 - F_L(\xi)]}{1 - F_L(\xi)}$$

$$+ \frac{F_L(\xi) - (1/2)\int_0^\xi (1 - F_T)\, dF_L}{1 - (1/2)\int_0^\xi (1 - F_T)\, dF_L} \times \frac{\int_0^\xi F_T\, dF_L}{F_L(\xi) - (1/2)\int_0^\xi (1 - F_T)\, dF_L} \quad . \tag{59}$$

We assume $F_L(\xi) < 1$, which must be true if there are any individuals left in observation after ξ. Expression (59) then reduces to a weighted mean of the form

$$A \cdot F_T(\xi) + (1 - A) \cdot \frac{\int_0^\xi F_T\, dF_L}{F_L(\xi) - (1/2)\int_0^\xi (1 - F_T)\, dF_L} \tag{60}$$

with $0 < A \leqslant 1$. Since $p^*_{0,\xi}$ is consistent if $q^*_{0,\xi}$ tends in probability to $F_T(\xi)$, i.e., when the expression (60) is equal to $F_T(\xi)$, we obtain

$$F_T(\xi) = \frac{\int_0^\xi F_T\, dF_L}{F_L(\xi) - (1/2)\int_0^\xi (1 - F_T)\, dF_L} \qquad \text{for all } \xi > 0 \tag{61}$$

as a necessary and sufficient condition for $p^*_{0,\xi}$ being consistent for every I_ξ. Abbreviating

$$G(\xi) = \frac{\int_0^\xi F_T\, dF_L}{F_L(\xi)} \quad , \tag{62}$$

we rewrite equation (61) as

$$F_T(\xi) = \frac{2G(\xi)}{1 + G(\xi)} \quad . \tag{63}$$

From equation (62) one verifies that

$$\frac{G'}{G}\,\frac{1+G}{1-G} = \frac{F_L'}{F_L} \quad ,$$

a differential equation which has the solution

$$\frac{G}{(1-G)^2} = cF_L \qquad \text{for } c > 0. \tag{64}$$

Solving equation (64) for G in terms of F_L and substituting in equation (63) one has

$$F_T(\xi) = 1 - \frac{1}{\sqrt{1 + cF_L(\xi)}} \tag{65}$$

We see, therefore, that relationship (65) between the probability distribution functions F_T and F_L is necessary and sufficient for the statistic (53) being a consistent estimate of $p_{a,b}$ for all intervals $(a, b]$.

In practical situations, it appears quite unrealistic to assume that T and L are not only independent but even have probability distributions related by the very stringent condition (65). One arrives, therefore, at the disappointing conclusion that the "adjusted observed" estimates of the probabilities $p_{a,b}$ are in general not consistent. It would be interesting to obtain bounds on the asymptotic bias of these estimates, but such bounds do not appear to be known.

3.3. LIFE-TABLE ESTIMATES: PROPORTIONAL HAZARD RATES

3.3.1. Definitions and Assumptions

We consider risks $R_1, R_2, \cdots, R_r, \cdots, R_k$. The time axis is divided by points

$$0 < u_1 < u_2 < \cdots < u_j < \cdots$$

chosen once for all (e.g., the last days of consecutive calendar years) into fixed time intervals $I_j = (u_{j-1}, u_j]$ for $j = 1, 2, \ldots$. One starts out with a cohort of N individuals, and observes the numbers

$$\delta_{rj} = \text{number of failures due to } R_r \text{ in } I_j. \tag{66}$$

We make the assumption that the net lives $X_1, X_2, \cdots, X_k$, corresponding to the k risks, are independent random variables, and wish to estimate the "net probabilities"

$$q_{rj} = P(X_r \leqslant u_j \mid X_r > u_{j-1}). \tag{67}$$

Let $\lambda_r(t)$ denote the hazard rate for the net life $X_r, r = 1, 2, \cdots, k$ and $\lambda(t)$ the hazard rate for the actual life $W = \min(X_1, \cdots, X_k)$. Since the X_r are independent, we have according to equation (19)

$$\lambda(t) = \sum_{r=1}^{k} \lambda_r(t). \tag{68}$$

The probability that in the presence of all competing risks an individual, alive at u_{j-1}, will survive to u_j can be written

$$p_j = P(W > u_j \mid W > u_{j-1}) = P(W > u_j)/P(W > u_{j-1})$$

$$= \exp\left[- \int_{u_{j-1}}^{u_j} \lambda(s)\, ds \right] = 1 - q_j. \tag{69}$$

The "crude" probability that an individual, alive at u_{j-1}, will fail of risk R_r during I_j is

$$Q_{rj} = P(X_r \in I_j \text{ and } X_r < X_s \text{ for } s \neq r \mid W > u_{j-1})$$

and can be written

$$Q_{rj} = \int_{u_{j-1}}^{u_j} \exp\left[- \int_{u_{j-1}}^{t} \lambda(s)\, ds \right] \lambda_r(t)\, dt. \tag{70}$$

We now make, in addition to the assumption of independent net lives, the following *assumption of proportional hazard rates* (PHR): For each interval I_j there are constants $c_{1j}, c_{2j}, \cdots, c_{rj}, \cdots, c_{kj}$ such that

$$\frac{\lambda_r(t)}{\lambda(t)} = c_{rj} \qquad \text{for } t \in I_j. \tag{71}$$

Under this assumption, the crude probabilities of equation (70) can be expressed as

$$Q_{rj} = c_{rj} \int_{u_{j-1}}^{u_j} \exp\left[- \int_{u_{j-1}}^{t} \lambda(s)\, ds \right] \lambda(t)\, dt$$

$$= c_{rj} \left\{ 1 - \exp\left[- \int_{u_{j-1}}^{u_j} \lambda(s)\, ds \right] \right\},$$

hence

$$Q_{rj} = c_{rj} q_j \qquad \text{for } r = 1, \cdots, k; j = 1, 2, \cdots, \tag{72}$$

where q_j is defined by equation (69) as the probability that an individual, alive at u_{j-1}, will fail in I_j.

From equations (71) and (72) one has

$$\frac{\lambda_r(t)}{\lambda(t)} = \frac{Q_{rj}}{q_j} = c_{rj} \qquad \text{for } t \text{ in } I_j. \tag{73}$$

Summing over r and using equation (68) one sees that

$$\sum_{r=1}^{k} Q_{rj} = q_j. \tag{74}$$

Equation (74) follows directly from the definition of the Q_{rj} and does not require the PHR assumption. Using the PHR assumption, the net probability q_{rj} defined by equation (67) can be written

$$\begin{aligned} q_{rj} &= 1 - \exp\left[- \int_{u_{j-1}}^{u_j} \lambda_r(s)\, ds \right] \\ &= 1 - \exp\left[- c_{rj} \int_{u_{j-1}}^{u_j} \lambda(s)\, ds \right] \\ &= 1 - p_j^{c_{rj}} \end{aligned}$$

and by equation (72)

$$q_{rj} = 1 - p_j^{Q_{rj}/q_j}, \tag{75}$$

which expresses the net probability q_{rj} in terms of the probabilities p_j and $q_j = 1 - p_j$, and the crude probability Q_{rj}.

3.3.2. Observable Random Variables: Maximum Likelihood Estimates of Crude and Net Probabilities

We assume that, as is most often the case in life-table-type studies, the records contain only the numbers

$$\begin{aligned} l_j &= \text{number of those alive immediately after } u_{j-1} \\ d_{rj} &= \text{number of those failing in } I_j \text{ due to } R_r \end{aligned} \qquad \text{for } j = 1, 2, \ldots; r = 1, \ldots, k. \tag{76}$$

Clearly,

$$l_j = \sum_{r=1}^{k} d_{rj} + l_{j+1}. \tag{77}$$

Given that an individual survived u_{j-1}, the probabilities that it will fail due to $R_1, \ldots, R_r, \ldots, R_k$ are, respectively, $Q_{1j}, \ldots, Q_{rj}, \ldots, Q_{kj}$, and the probability that it will survive beyond u_j is p_j. Therefore, the conditional joint probability distribution of $d_{1j}, \ldots, d_{rj}, \ldots, d_{kj}, l_{j+1}$, given l_j, is the multinomial expression

$$\frac{l_j!}{d_{1j}! \cdots d_{kj}! l_{j+1}!}\, Q_{1j}^{d_{1j}} \cdots Q_{kj}^{d_{kj}} p_j^{l_{j+1}}. \tag{78}$$

When for an interval I_j the observed frequencies l_j, $d_{1j}, \ldots, d_{kj}$ and l_{j+1} are available from observation, then it is well known that the maximum likelihood estimates for the probabilities appearing in expression (78) are the relative frequencies

$$\hat{Q}_{rj} = \frac{d_{rj}}{d_j} \qquad \text{for } r = 1, 2, \ldots, k$$
$$\hat{p}_j = \frac{l_{j+1}}{l_j}, \tag{79}$$

and expressions for variances and covariances can be explicitly obtained (p. 253[10]).

Thus far, the arguments have been straightforward and fairly simple, but they yielded only the estimates (79) of the crude probabilities. To estimate the net probabilities q_{rj}, Chiang recommends that the values (79) can be substituted in equation (75), so that one obtains the estimates

$$q_{rj} = 1 - \left(\frac{l_{j+1}}{l_j}\right)^{d_{rj}/(l_j - l_{j+1})}$$
$$= 1 - \left(\frac{l_{j+1}}{l_j}\right)^{d_{rj}/d_j} \tag{80}$$

where $d_j = l_j - l_{j+1}$ = total number of failures in I_j. Since in most practical situations one is mainly interested in the net probabilities, the question arises whether the estimates (80) have the usual desirable properties and in particular whether they are consistent.

3.3.3. Conditions for Consistency of Chiang's Estimator of the Net Survival Probabilities

As in section 3.2.2, we consider the case of $k = 2$ competing risks, call the corresponding lives T = time to failure and L = time to loss from observation, assume that T and L are independent random variables, and denote their survival probabilities by $\bar{F}_T(t)$ and $\bar{F}_L(t)$ and their probability densities by $f_T(t)$ and $f_L(t)$. Again, without loss of generality, we fix our attention on an interval $I_\xi = (0, \xi]$, $\xi > 0$, and for given initial size N of the cohort use the estimator (80).

Using the events defined by equations (55) and the random variables in (56) we rewrite (80) in the form

$$\hat{q}_{T,\xi} = 1 - \left(\frac{N - D - M}{N}\right)^{D/(D+M)}$$
$$= 1 - \left(1 - \frac{D_1}{N} - \frac{D_2}{N} - \frac{M}{N}\right)^{(D_1/N + D_2/N)/(D_1/N + D_2/N + M/N)} \tag{81}$$

Again, as in section 3.2.2 we conclude that with $N \to \infty$ each of the quotients in this expression tends in probability to the corresponding probability in (57), hence $\hat{q}_{T,\xi}$ tends in probability to

$$q^* = 1 - [1 - P(\delta_1) - P(\delta_2) - P(\mu)]^{[P(\delta_1)+P(\delta_2)]/[P(\delta_1)+P(\delta_2)+P(\mu)]} \tag{82}$$

As can be seen from figure 6, the sets $\delta = \delta_1 \cup \delta_2$ and μ correspond to the symmetric, observable events "failure before ξ and before loss" and "loss before ξ and before failure," and their probabilities $P(\delta), P(\mu)$ satisfy the equation

$$P(\delta) + P(\mu) + P(L > \xi, T > \xi) = P(\delta) + P(\mu) + \bar{F}_L(\xi)\bar{F}_T(\xi) = 1,$$

so that equation (82) can be written

$$q^* = 1 - [\bar{F}_L(\xi)\bar{F}_T(\xi)]^{[P(\delta)]/[1-\bar{F}_L(\xi)\bar{F}_T(\xi)]}. \tag{83}$$

Therefore, for $\hat{q}_{T,\xi}$ to be consistent, it is necessary and sufficient that

$$[\bar{F}_L(\xi) \cdot \bar{F}_T(\xi)]^{[P(\delta)]/[1-\bar{F}_L(\xi)\bar{F}_T(\xi)]} = \bar{F}_T(\xi)$$

which reduces to

$$[\bar{F}_L(\xi)]^{P(\delta)} = [\bar{F}_T(\xi)]^{P(\mu)}. \tag{84}$$

If the PHR assumption (71) is satisfied, which in our case of two independent competing risks means

$$\lambda_T(s) = c\lambda(s), \lambda_L(s) = (1 - c)\lambda(s),$$

then

$$\bar{F}_T(u) = \exp\left[-c\int_0^u \lambda(s)\,ds\right],$$

$$\bar{F}_L(u) = \exp\left[-(1-c)\int_0^u \lambda(s)\,ds\right]$$

and

$$P(\mu) = -\int_{u=0}^{\xi} \bar{F}_T(u)\,d\bar{F}_L(u)$$

$$= (1-c)\left\{1 - \exp\left[-\int_0^{\xi} \lambda(s)\,ds\right]\right\}$$

$$P(\delta) = -\int_0^{\xi} \bar{F}_L(u)\,d\bar{F}_T(u)$$

$$= c\left\{1 - \exp\left[-\int_0^{\xi} \lambda(s)\,ds\right]\right\}$$

and one verifies that condition (84) is satisfied. We conclude that *under the PHR assumption, estimate (80)* proposed by Chiang *is consistent.*

3.3.4. Examples of Competing Risks With Proportional Hazard Rates

We have seen that for data of the life-table type, under the assumptions of independence and proportional hazard rates, Chiang's estimate (80) is consistent. Leaving aside the question of how realistic these assumptions may be in specific practical situations, one may still wish to see nontrivial examples of independent competing risks which satisfy the PHR assumption. Following H. A. David in 1970,[27] we present a further discussion of the PHR property (71), which leads to the construction of such examples.

Writing equation (71) as

$$\lambda_r(t) = c_{rj}\lambda(t) \qquad \text{for } u_{j-1} < t \leqslant u_j \tag{85}$$

and integrating from u_{j-1} to $t(\leqslant u_j)$, one obtains according to equation (6) the relationships

$$\log \frac{\bar{F}_{X_r}(t)}{\bar{F}_{X_r}(u_{j-1})} = c_{rj} \log \frac{\bar{F}_W(t)}{\bar{F}_W(u_{j-1})}$$

where $W = \min(X_1, \ldots, X_k)$. This shows that

$$\left[\frac{\bar{F}_{X_r}(t)}{\bar{F}_{X_r}(u_{j-1})}\right]^{1/c_{rj}} \tag{86}$$

is independent of r for fixed j and $u_{j-1} < t \leqslant u_j$.

Consider now the more restrictive special case of the PHR assumption (85)

$$\lambda_r(t) = c_r\lambda(t) \qquad \text{for } 0 < t < \infty; r = 1, \ldots, k. \tag{87}$$

Then expression (86) can be written

$$[\bar{F}_{X_r}(t)]^{1/c_r} = [\bar{F}_{X_1}(t)]^{1/c_1} \qquad \text{for } 0 < t < \infty; r = 2, \ldots, k$$

or

$$[\bar{F}_{X_r}(t)]^{c_1/c_r} = \bar{F}_{X_1}(t) \qquad \text{for } 0 < t < \infty; r = 2, \ldots, k.$$

If the $c_1, \ldots, c_k$ are so chosen that

$$p_r = \frac{c_1}{c_r} \qquad \text{for } r = 2, \ldots, k$$

are positive integers, then one has

$$\bar{F}_{X_1}(t) = [\bar{F}_{X_r}(t)]^{p_r} \qquad \text{for } 0 < t < \infty; r = 2, \ldots, k, \tag{88}$$

hence X_1 can be represented, for each $r = 2, \ldots, k$, as the minimum of p_r independent identically distributed random variables with survival function $\bar{F}_{X_r}(t)$.

To obtain a nice example of net lives that satisfy relation (88), David considers among others the family of Weibull distributions

$$P(X > t; \alpha, \beta) = \bar{F}(t; \alpha, \beta) = \exp\,[-\alpha t^{\beta}] \tag{89}$$

which has the well known property that the minimum of independent identically distributed random variables of this family has again a Weibull distribution (his discussion goes further, to consider all three known classes of extreme-value distributions of the minimum). Choosing $c_1, c_2, \ldots, c_k$ so that for $r = 2, \ldots, k$ the ratios p_r are positive integers, and assuming for X_1 a Weibull distribution (89), one obtains an example of independent net lives, each with a Weibull distribution, which satisfy condition (87), i.e., have proportional hazard rates on the entire time axis. For suggestions of further examples the reader may wish to consult David's 1970 paper.[27]

4. ESTIMATION FOR PARAMETRIC MODELS

4.1. THE LIKELIHOOD FUNCTION: GENERAL CASE

We shall now assume that the joint probability distribution of the net lives belongs to a specified parametric family. Let the joint probability density of these net lives $X_1, \ldots, X_k$ be

$$f_{\underline{X}}(\underline{x}; \underline{\theta}) = f(x_1, \ldots, x_k; \theta_1 \ldots, \theta_m). \tag{90}$$

For fixed j, let

$$\pi_j = P(\min_{i=1,\ldots,k} X_i = X_j) \qquad \text{for } j = 1, \ldots, k, \tag{91}$$

be the probability that failure due to risk R_j is observed. Writing in equation (11) $f_{\underline{X}}(\underline{x}; \underline{\theta})$ instead of $f(x_1, \ldots, x_k)$, one obtains

$$\pi_j = \int_{x_j=0}^{\infty} \int_{x_1=x_j}^{\infty} \cdots \int_{x_{j-1}=x_j}^{\infty} \int_{x_{J+1}=x_j}^{\infty} \cdots \int_{x_k=x_j}^{\infty} f_{\underline{X}}(\underline{x}; \underline{\theta}) \prod_{i \neq j} dx_i\, dx_j \tag{92}$$

and the probability density of the "crude" life Y_j, as defined in section 2.4, can be written

$$f_{Y_j}(y_j; \underline{\theta}) = \frac{1}{\pi_j} \int_{y_j}^{\infty} \cdots \int_{y_j}^{\infty} f(x_1, \ldots, x_{j-1}, Y_j, x_{j+1}, \ldots, x_k; \theta_1, \ldots \theta_m) \prod_{i \neq j} dx_i. \tag{93}$$

When n individuals are observed to failure, and n_j of them fail due to R_j, $j = 1, \ldots, k$, then the joint probability density of the observed (crude) life lengths Y_{jr}, $r = 1, \ldots, n_j$, conditional on the n_j, is

$$\prod_{j=1}^{k} 1/\pi_j^{n_j} \prod_{r=1}^{n_j} \int_{y_{jr}}^{\infty} \cdots \int_{y_{jr}}^{\infty} f(x_1, \ldots, x_{j-1}, y_{jr}, x_{j+1}, \ldots, x_k; \underline{\theta}) \prod_{i \neq j} dx_i.$$

Since the n_j have a multinomial probability distribution

$$g(n_1, \ldots, n_k) = \left(n! \Big/ \prod_{j=1}^{k} n_j! \right) \left(\prod_{j=1}^{k} \pi_j^{n_j} \right),$$

the likelihood function of the observed (crude) life lengths is

$$L = L(\ldots, y_{jr}, \ldots; \theta_1, \ldots, \theta_m) = \left(n! \Big/ \prod_{j=1}^{k} n_j! \right) \prod_{j=1}^{k} \prod_{r=1}^{n_j} \int_{y_{jr}}^{\infty} \cdots \int_{y_{jr}}^{\infty}$$

$$f(x_1, \ldots, x_{j-1}, y_{jr}, x_{j+1}, \ldots, x_k; \theta_1, \ldots, \theta_m) \prod_{i \neq j} d_{x_i}. \tag{94}$$

This expression, given in 1971 by Moeschberger and David,[28] does not require the assumption that the net lives are independent, but only the assumption of a specified parametric form of the joint probability density (equation 90) of the net lives. Whenever such a parametric family is specified and the values of the crude lives y_{jr} are available, the likelihood function (94) may be used to obtain maximum likelihood estimates of the parameters $\theta_1, \ldots \theta_m$ by the usual procedure of computing the partial derivatives

$$\frac{\partial L}{\partial \theta_s} \left(\text{or } \frac{\partial \ln L}{\partial \theta_s} \right),$$

equating them to zero, and solving the resulting system of m equations for $\theta_1, \ldots, \theta_m$. However, even if the density (90) is of a reasonably simple form, one would expect in carrying out these steps to encounter difficulties which, at best, may be overcome by using computers.

4.2. THE CASE OF INDEPENDENT NET LIVES

If one assumes that the net lives $X_1, \ldots, X_k$ are independent, with probability densities $f_j(x_j; \underline{\theta})$ and survival functions $\bar{F}_j(x_j; \underline{\theta})$, then equation (93) becomes

$$f_{Y_j}(y_j; \underline{\theta}) = \frac{f_j(y_j; \underline{\theta})}{\pi_j \bar{F}_j(y_j; \theta)} \prod_{l=1}^{k} \bar{F}_l(y_l; \underline{\theta}) \tag{95}$$

and the likelihood function (94) takes the form

$$L = \left(n! \Big/ \prod_{j=1}^{k} n_j! \right) \prod_{j=1}^{k} \prod_{r=1}^{n_j} [f_j(y_{jr}; \underline{\theta})] / [\bar{F}_j(y_{jr}; \underline{\theta})] \prod_{l=1}^{k} \bar{F}_l(y_{jr}; \underline{\theta}). \tag{96}$$

This expression is much less complicated than (94). Moreover, for simple parametric models, it lends itself to manageable treatment for obtaining maximum likelihood estimates of the parameters $\underline{\theta} = (\theta_1, \ldots, \theta_m)$.

4.3. COMMENTS ON PARAMETRIC FAMILIES OF MULTIVARIATE LIFE DISTRIBUTIONS

The maximum likelihood estimation procedure under the assumption of independent net lives, as just described, has been actually used. The reader may find technical details as well as completely worked out numerical examples in the 1971 paper by Herman and Patell[29] or in Moeschberger and David[28] for the cases when the net lives are assumed to have independent exponential or Weibull distributions. Moeschberger and David consider also the multivariate exponential distribution introduced in 1967 by Marshall and Olkin,[30] which deals with *dependent* net lives, and suggest a way of proceeding in this case.[d]

The Marshall-Olkin multivariate exponential distribution is one of the few multivariate life distributions admitting dependence. A multivariate Weibull distribution has been proposed by several authors, and a rather general method for obtaining new multivariate life distributions was outlined by Lee and Thompson.[31] To this writer's knowledge, little work has been done on the use of such distributions. It would seem desirable to construct more parametric families of multivariate life distributions, justify their use either by deriving them from some plausible assumptions, or by at least showing that they agree reasonably well with empirical data, and to develop in detail the maximum likelihood estimation techniques for their parameters.

4.4. CONCOMITANT VARIABLES

In some practical situations, the available empirical data contain more information than just the observed crude lives (i.e., actual lives and the risk recorded at failure), and it appears desirable to use this additional information. An important example is the situation when values of "concomitant variables" have been observed. Considerable advances have been made recently in developing techniques for using concomitant variables and, while a systematic presentation of this area of research would require a separate monograph, it may be useful to give here an example of the concepts involved and mention some of the pertinent papers.

In 1965 Feigl and Zelen[32] considered the study of data on leukemia patients which, for each patient, contain the observed time from diagnosis to death (life length) and the white blood cell count at diagnosis (concomitant variable). They assumed that the probability distribution of life length T at diagnosis depends on the white cell count x in such a way that the probability density of T is

$$F(t) = \begin{cases} \exp(-\lambda t) & \text{for } t \geqslant 0 \\ 0 & \text{for } t < 0 \end{cases}$$

where

$$E(T) = \frac{1}{\lambda} = a + bx.$$

The problem was to estimate the parameters a, b which determine the dependence of T on x; knowing these parameters would clearly help in estimating the life expectancy of a patient for whom a value of

[d]The Marshall-Olkin distribution is also discussed in Langberg, Proschan, and Quinzi's 1978 paper.[20]

x was observed. Feigl and Zelen derived the maximum likelihood procedure for this problem, computed the asymptotic variance-covariance matrix, and applied the theory to numerical data. In 1966 Zippin and Armitage[33] extended the Feigl-Zelen model to the case in which not all patients had died at the time when the study was concluded, thus introducing a competing risk.

The important paper of 1972 by Cox[34] considered a very general model dealing with both censoring (competing risks) and concomitant variables. This paper was followed by a number of studies, such as Kalbfleisch and Prentice[35] in 1973, Lagakos[36] in 1976, and Holford[37] also in 1976.

4.5. ESTIMATION OF THE RATIO OF HAZARD FUNCTIONS

As an example of a problem which requires estimating a parameter from data with competing risks, we refer to a 1975 paper by Crowley[24] in which he discussed the following situation.

Survival data were obtained independently for two different populations, each affected by competing risks. There is reason to assume that the two net life distributions are of the same type and that the difference between them is due only to the fact that their hazard rates differ by a constant factor. How can one estimate that factor?

Let X_1 denote the net life in the first population, and T_1 the competing life, so that only the actual life

$$W_1 = \min(X_1, T_1)$$

and an indicator variable

$$\delta_1 = \begin{cases} 1 & \text{when } X_1 \leqslant T_1 \\ 0 & \text{when } X_1 > T_1 \end{cases}$$

can be observed, and let X_2, T_2, W_2, δ_2 be similarly defined for the second population. We shall say that W_1 is an uncensored value when $\delta_1 = 1$, and a censored value when $\delta_1 = 0$, and use similar terms for W_2.

Available are the samples of (W_1, δ_1)

$$(W_{1j}, \delta_{1j}), j = 1, 2, \ldots, n_1 \tag{97}$$

and of (W_2, δ_2)

$$(W_{2k}, \delta_{2k}), k = 1, 2, \ldots, n_2. \tag{98}$$

We assume that (1) X_1, T_1, X_2, and T_2 are independent and (2) the hazard rates of X_1 and X_2 differ only by a constant factor, i.e.,

$$\lambda_{x_2}(t) = \theta \lambda_{x_1}(t) \qquad \text{for } t \geqslant 0. \tag{99}$$

Assumption (99) is equivalent to the relationship between survival functions

$$\bar{F}_{X_2}(t) = [\bar{F}_{X_1}(t)]^{\theta}. \tag{100}$$

If these assumptions are satisfied, then $\theta > 1$ implies that X_2 is stochastically smaller than X_1.

Crowley considered several possibilities of presenting the data (individual sample values are available, or data are grouped in time intervals, etc.) and discussed several estimators. We shall describe one of them, for the case when the exact individual values (97) and (98) are available. Let

$$n_{11} = \sum_{j=1}^{n_1} \delta_{1j} = \text{number of uncensored values of } W_1$$

$$n_{12} = \sum_{j=1}^{n_1} (1 - \delta_{1j}) = \text{number of censored values of } W_1$$

so that $n_{11} + n_{12} = n_1$ and, similarly,

$$n_{21} = \sum_{k=1}^{n_2} \delta_{2k} \qquad n_{22} = \sum_{k=1}^{n_2} (1 - \delta_{2k}),$$

with $n_{21} + n_{22} = n_2$. We denote the uncensored values of W_1 by $W_{1,1}, W_{1,2}, \ldots, W_{1,n_{11}}$; the censored values of W_1 by $W_{1,n_{11}+1}, W_{1,n_{11}+2}, \ldots, W_{1,n_{11}+n_{12}}$; the uncensored values of W_2 by $W_{2,1}, W_{2,2}, \ldots, W_{2,n_{21}}$; and the censored values of W_2 by $W_{2,n_{21}+1}, \ldots, W_{2,n_{21}+n_{22}}$.

The likelihood function (96) becomes now

$$\frac{n_1!}{n_{11}!n_{12}!} \prod_{r=1}^{n_{11}} f_{X_1}(W_{1r})\bar{F}_{T_1}(W_{1r}) \prod_{r=n_{11}+1}^{n_{11}+n_{12}} f_{T_1}(W_{1r})\bar{F}_{X_1}(W_{1r})$$

$$\times \frac{n_2!}{n_{21}!n_{22}!} \prod_{s=1}^{n_{21}} f_{X_2}(W_{2s})\bar{F}_{T_2}(W_{2s}) \prod_{s=n_{21}+1}^{n_{21}+n_{22}} f_{T_2}(W_{2s})\bar{F}_{X_2}(W_{2s}). \tag{101}$$

To carry out the procedure leading to a solution of the likelihood equation for θ, one must assume a specific functional expression for $\bar{F}_{X_1}(t)$. Following Crowley, we illustrate this procedure by choosing

$$\bar{F}_{X_1}(t) = \exp(-\lambda t) \tag{102}$$

hence

$$\bar{F}_{X_2}(t) = \exp(-\theta\lambda t). \tag{103}$$

Writing expression (101) for these specific survival functions and differentiating $\log L$ with respect to λ and to θ, one arrives at the likelihood equations

$$\frac{n_{11} + n_{21}}{\lambda} - \sum_{r=1}^{n_1} W_{1r} - \theta \sum_{s=1}^{n_2} W_{2s} = 0$$

$$\frac{n_{21}}{\theta} - \lambda \sum_{s=1}^{n_2} W_{2s} = 0,$$

and the estimators for λ and θ obtained by solving these equations are

$$\hat{\lambda} = n_{11} \Big/ \sum_{r=1}^{n_1} W_{1r}$$

$$\hat{\theta} = (n_{21}/n_{11}) \sum_{r=1}^{n_1} W_{1r} \Big/ \sum_{s=1}^{n_2} W_{2s} \quad . \tag{104}$$

Estimate (104) is consistent, since it can be written:

$$\hat{\theta} = \left(\frac{n_{21}}{n_2} \Big/ \frac{n_{11}}{n_1} \right) \left(\frac{\sum_{r=1}^{n_1} W_{1r}}{n_1} \Big/ \frac{\sum_{s=1}^{n_2} W_{2s}}{n_2} \right)$$

and this tends in probability to

$$\frac{P(X_2 \leqslant T_2)}{P(X_1 \leqslant T_1)} \frac{E(W_1)}{E(W_2)} = \frac{\int_0^\infty \bar{F}_{T_2}(t)\, d\bar{F}_{X_2}(t)}{\int_0^\infty \bar{F}_{T_1}(t)\, d\bar{F}_{X_1}(t)} \cdot \frac{\int_0^\infty t\, d[\bar{F}_{X_1}(t)\bar{F}_{T_1}(t)]}{\int_0^\infty t\, d[\bar{F}_{X_2}(t)\bar{F}_{T_2}(t)]}$$

$$= \frac{\int_0^\infty \bar{F}_{T_2}(t)\, d[\exp(-\theta\lambda t)]}{\int_0^\infty \bar{F}_{T_1}(t)\, d[\exp(-\lambda t)]} \cdot \frac{\int_0^\infty t\, d[\exp(-\lambda t)\bar{F}_{T_1}(t)]}{\int_0^\infty t\, d[\exp(-\theta\lambda t)\bar{F}_{T_2}(t)]}$$

$$= \theta.$$

This argument makes use of equations (102) and (103), i.e., of the assumption of exponential net lives. In general, consistency and asymptotic normality of the estimate obtained from the likelihood function (101) will follow from known properties of maximum likelihood estimates.

5. TESTS OF HYPOTHESES

5.1. FORMULATION OF A PROBLEM

In clinical studies dealing with the comparison of two treatments one often wishes to test the hypothesis that these treatments are equally effective, i.e., that lives (times from beginning of treatment to the onset of some condition or to death) have the same probability distribution under both treatments, against some alternative indicating that one of the treatments is superior. When there are

no competing risks, a number of statistical tests are readily available for this purpose, such as the Wilcoxon-Mann-Whitney test. The case when patients under each treatment are exposed to a competing risk, such as loss from observation (censoring on the right), was first considered in 1965 by Gehan.[38] Similar techniques had been proposed in 1962 in an unpublished thesis by Gilbert.[39] In 1965 Gehan[40] extended the test procedure to doubly censored data. A further study of properties of the Gehan statistic was made in 1967 by Efron[41] who proposed additional test statistics. Breslow[42] studied (in 1970) the more general problem of comparing $k \geqslant 2$ treatments and testing the hypothesis that they are equally effective, when each of the available k samples is subject to competing risks and the competing net lives have possibly different probability distributions.

5.2. GEHAN'S STATISTIC

In presenting the contents of Gehan,[38] we shall follow some of the arguments used by Efron.[41]

Let X and U be independent lives corresponding to competing risks, with survival functions $\bar{F}_X(t)$ and $\bar{F}_U(t)$; and let Y and V be another pair of independent lives corresponding to competing risks, with survival functions $\bar{F}_Y(t)$, $\bar{F}_V(t)$. To avoid complications due to the possiblity of ties, we shall assume that all survival functions are continuous. Available are independent random samples

$$X_1, X_2, \ldots, X_m \text{ of } X, \qquad \text{hence with survival function } \bar{F}_X$$

$$U_1, U_2, \ldots, U_m \text{ of } U, \qquad \text{hence with survival function } \bar{F}_U$$

$$Y_1, Y_2, \ldots, Y_n \text{ of } Y, \qquad \text{hence with survival function } \bar{F}_Y$$

$$V_1, V_2, \ldots, V_n \text{ of } V, \qquad \text{hence with survival function } \bar{F}_V.$$

Since we deal with competing risks, we can observe only

$$X_i^* = \min(X_i, U_i) \tag{105}$$

and

$$\delta_i = \begin{cases} 1 \text{ if } X_i^* = X_i \\ 0 \text{ if } X_i^* = U_i \end{cases} \tag{106}$$

for $i = 1, 2, \ldots, m$, and similarly

$$Y_j^* = \min(Y_j, V_j) \tag{107}$$

and

$$\epsilon_j = \begin{cases} 1 \text{ if } Y_j^* = Y_j \\ 0 \text{ if } Y_j^* = V_j \end{cases} \qquad \text{for } j = 1, 2, \ldots, n. \tag{108}$$

Clearly, the random variables X_i^*, Y_j^*, are mutually independent, and their survival functions are

$$\begin{aligned} \bar{F}_{X_i^*}(s) &= \bar{F}_X(s)\bar{F}_U(s) \qquad \text{for } i = 1, 2, \ldots, m \\ \bar{F}_{Y_j^*}(s) &= \bar{F}_Y(s)\bar{F}_V(s) \qquad \text{for } j = 1, 2, \ldots, n, \end{aligned} \tag{109}$$

while the δ_i and ϵ_j are mutually independent Bernoulli variables with probabilities

$$\begin{aligned} P(\delta_i = 1) &= P(X_i \leqslant U_i) \\ &= -\int_0^\infty \bar{F}_U(s)\, d\bar{F}_X(s) \\ P(\epsilon_j = 1) &= P(Y_j \leqslant V_j) \\ &= -\int_0^\infty \bar{F}_V(s)\, d\bar{F}_Y(s). \end{aligned} \tag{110}$$

In general, X_i^* and δ_i are not independent, and neither are Y_j^*, ϵ_j.

The following notation used by Efron simplifies writing some of the arguments: For two independent random variables S, T, we shall write

$$\begin{aligned} P(F_s \geqslant \bar{F}_t) &= -\int_{-\infty}^\infty \bar{F}_S(z)\, d\bar{F}_T(z) \\ &= P(S \geqslant T). \end{aligned} \tag{111}$$

For every observed four values X_i^*, δ_i, Y_j^*, ϵ_j, we define a scoring function $Q(X_i^*, \delta_i, Y_j^*, \epsilon_j)$ by assigning to it the value 1 when the four values establish unequivocally that $X_i \geqslant Y_j$, the value 0 when the four values establish unequivocally that $X_i < Y_j$, and the value 1/2 in all other cases. More formally

$$Q(X_i^*, \delta_i, Y_j^*, \epsilon_j) = \begin{cases} 1 & \text{when } \epsilon_j = 1 \text{ and } X_i^* \geqslant Y_j^* \\ 0 & \text{when } \delta_i = 1 \text{ and } X_i^* < Y_j^* \\ 1/2 & \text{otherwise.} \end{cases} \tag{112}$$

One verifies that this definition of the scoring function is equivalent with

$$Q(X_i^*, \delta_i, Y_j^*, \epsilon_j) = \begin{cases} 1 & \text{when } Y_j \leqslant \min(X_i, U_i, V_j) \\ 0 & \text{when } X_i < \min(Y_j, V_j, U_i) \\ 1/2 & \text{when } U_i < \min(X_i, Y_j, V_j) \\ 1/2 & \text{when } V_j < \min(X_i, Y_j, U_i). \end{cases} \tag{113}$$

Gehan's test statistic is defined in terms of this scoring function by the expression

$$W = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} Q(X_i^*, \delta_i, Y_j^*, \epsilon_j). \tag{114}$$

Using equation (113) and the notation introduced in equation (111), we have

$$P[Q(X_i^*, \delta_i, Y_j^*, \epsilon_j) = 1] = P(\bar{F}_Y \leqslant \bar{F}_X \bar{F}_U \bar{F}_V)$$

$$P[Q(X_i^*, \delta_i, Y_j^*, \epsilon_j) = 0] = P(\bar{F}_X < \bar{F}_Y \bar{F}_V \bar{F}_U)$$

$$P[Q(X_i^*, \delta_i, Y_j^*, \epsilon_j) = 1/2] = P(\bar{F}_U < \bar{F}_X \bar{F}_Y \bar{F}_V) + P(\bar{F}_V < \bar{F}_X \bar{F}_Y \bar{F}_U).$$

Hence the expectation of W can be written

$$E(W) = P(\bar{F}_Y \leqslant \bar{F}_X \bar{F}_U \bar{F}_V) + (1/2) \cdot [P(\bar{F}_U < \bar{F}_X \bar{F}_Y \bar{F}_V) + P(\bar{F}_V < \bar{F}_X \bar{F}_Y \bar{F}_U)]. \tag{115}$$

Under the null hypothesis (H_0)

$$\bar{F}_X(t) = \bar{F}_Y(t) \qquad \text{for all } t \geqslant 0,$$

one has

$$P(\bar{F}_Y \leqslant \bar{F}_X \bar{F}_U \bar{F}_V) = P(\bar{F}_X \leqslant \bar{F}_Y \bar{F}_U \bar{F}_V)$$

and equation (115) becomes

$$\begin{aligned} E(W) = (1/2)[&P(\bar{F}_X \leqslant \bar{F}_Y \bar{F}_U \bar{F}_V) + P(\bar{F}_Y \leqslant \bar{F}_X \bar{F}_U \bar{F}_V) \\ &+ P(\bar{F}_U < \bar{F}_X \bar{F}_Y \bar{F}_V) + P(\bar{F}_V < \bar{F}_X \bar{F}_Y \bar{F}_U)] = 1/2. \end{aligned} \tag{116}$$

It should be noted that equation (116) is true *for any two censoring distributions* F_U, F_Y.[c]

By a similar argument one obtains for the variance of W under the null hypothesis (H_0) the expression

$$\begin{aligned} \mathrm{Var}_{H_0}(W) = \frac{1}{12mn} &[3P(\bar{F}_X^2 < \bar{F}_U \bar{F}_Y) + (n-1)P(\bar{F}_X^3 < \bar{F}_U \bar{F}_V^2) \\ &+ (m-1)P(\bar{F}_X^3 < \bar{F}_U^2 \bar{F}_V)]. \end{aligned} \tag{117}$$

This expression for the variance of W was obtained in 1962 by Gilbert[39] and is quoted in 1967 in Efron's paper.[41]

---

[c] Gehan obtained equation (116) only under the additional assumption that the censoring distributions $\bar{F}_U(t)$ and $\bar{F}_V(t)$ are the same. In 1967 Mantel[43] observed that this assumption is not needed and he communicated this fact earlier to Efron for inclusion in his 1967 report.[41]

To test H_0, one would compute W from the available data and compare it with the expected value 1/2. To judge the significance of the difference $W - 1/2$, one has to keep in mind two facts pointed out by Efron:

The statistic W is not nonparametric, since, even under $H_0: \overline{F}_X = \overline{F}_Y$, its variance (117) depends on the relationship between $\overline{F}_X$, $\overline{F}_U$, and $\overline{F}_V$.

W is asymptotically nonparametric in the following sense: when $m \to \infty, n \to \infty$, so that $[m/(m+1)] \to \lambda$ where $0 < \lambda < 1$, then under H_0 the probability distribution function of

$$\sqrt{m+n}\,[W - (1/2)] \text{ tends to } N\left\{0, (1/12)\left[(1/\lambda)\sigma_1^2 + \frac{1}{1-\lambda}\,\sigma_2^2\right]\right\} \tag{118}$$

where

$$\begin{aligned} \sigma_1^2 &= P(\overline{F}_u \overline{F}_v^2 > \overline{F}_x^3), \\ \sigma_2^2 &= P(\overline{F}_u^2 \overline{F}_v > \overline{F}_x^3), \end{aligned} \tag{119}$$

that is, the random variable $\sqrt{m+n}\,(W - 1/2)$ is asymptotically normal with expectation 0 and variance

$$(1/12) \times \frac{1}{\lambda}\left(\sigma_1^2 + \frac{1}{1-\lambda}\,\sigma_2^2\right).$$

Furthermore, Efron suggests consistent estimates for σ_1^2 and σ_2^2, but does not elaborate on the way to compute them.

Under the additional assumption $\overline{F}_U = \overline{F}_V$, i.e., of the same probability distribution for the two censoring variables, Gehan[38] proposed in 1965 the use of W for a conditional test of H_0 of a combinatorial nature. Mantel[43] offers a simple combinatorial formulation of Gehan's conditional test statistic, and discusses its properties under the assumption $\overline{F}_U = \overline{F}_V$ and also under less restrictive assumptions.

5.3. TESTING FOR INDEPENDENCE WHEN DATA ARE CENSORED

An important problem of hypothesis testing, treated extensively in 1974 by Brown, Hollander, and Korwar,[44] may be stated as follows.

Consider the random variables X, Y with a bivariate life distribution, i.e., such that

$$P(X \geq 0, y \geq 0) = 1.$$

If a sample

$$(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n) \tag{120}$$

is available, then there are a number of well known tests of the hypothesis that X and Y are independent or, more generally, uncorrelated. A new problem arises when to one or both of the two lives X, Y corresponds a competing life (censoring variable), U in competition with X, and V in competition with

Y, so that one can only observe

$$(X_1^*, \delta_1, Y_1^*, \epsilon_1), \ldots, (X_1^*, \delta_i, Y_i^*, \epsilon_i), \ldots, (X_n^*, \delta_n, Y_n^*, \epsilon_n) \tag{121}$$

where X_i^*, δ_i, Y_i^*, ϵ_i are defined in equations (105)-(108).

Brown, Hollander, and Korwar[44] considered data from a heart transplant program, and explored hypotheses such as that sex (X) and survival time after transplant (Y) are independent, where Y is censored by

V = time to closing date of the research program.

To test the hypothesis of independence of X and Y when only Y is censored (a generalization to the case of X and Y both censored is straightforward), the authors modified Kendall's rank correlation statistic which for a complete sample (120) is defined as

$$S = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} b_{ij} \tag{122}$$

where

$$a_{ij} = \begin{cases} 1 & \text{if } X_i > X_j \\ 0 & \text{if } X_i = X_j \\ -1 & \text{if } X_i < X_j \end{cases} \tag{123}$$

and

$$b_{ij} = \begin{cases} 1 & \text{if } Y_i > Y_j \\ 0 & \text{if } Y_i = Y_j \\ -1 & \text{if } Y_i < Y_j. \end{cases} \tag{124}$$

When Y is censored, i.e., the sample (121) is of the form

$$(X_1, Y_1^*, \epsilon_1), \ldots, (X_i, Y_i^*, \epsilon_i), \ldots, (X_n, Y_n^*, \epsilon_n), \tag{125}$$

then the a_{ij} are still defined by equations (123), but equations (124) are modified to read

$$b'_{ij} = \begin{cases} 1 & \text{if } Y_i^* > Y_j^* \text{ and } \epsilon_j = 1, \text{ or } Y_i^* = Y_j^* \text{ and } \epsilon_i = 0 \\ -1 & \text{if } Y_i^* < Y_j^* \text{ and } \epsilon_i = 1, \text{ or } Y_i^* = Y_j^* \text{ and } \epsilon_j = 0 \\ 0 & \text{in all other cases} \end{cases} \tag{126}$$

and the test statistic of equation (122) is replaced by

$$S' = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b'_{ij}. \tag{127}$$

If X and Y are independent, then S' has a conditional probability distribution of a purely combinatorial nature, hence is nonparametric and can be used to test the hypothesis of independence of X and Y, provided the following additional assumption is satisfied:

Assumption A: When X and Y are independent, then X and (Y^*, ϵ) are independent.

If assumption A holds, critical values of S' can be calculated by a combinatorial argument when n is small. For large n, again under assumption A and when the hypothesis is true, the statistic S' is conditionally asymptotically normal with conditional expectation $E(S') = 0$ and a conditional variance that is given by Brown, Hollander, and Korwar[44] and is a function of the a_{ij} and b'_{ij}. A further statistic for testing independence of X, Y when one or both variables are censored, based on Kaplan-Meier extimates of the survival functions for the net lives X and Y, is proposed in their paper. The reader is referred to the original paper for a description of this test, as well as of still another "pseudoconditional" test.

5.4. LARGE-SAMPLE TESTS

In preceding sections we gave examples of tests of hypotheses about net lives which are observed in the presence of competing risks. These examples dealt with tests designed for specific problems, and the sampling distributions of the test statistics could in some cases be computed exactly for small sample sizes, and in all cases can be obtained for large samples from the asymptotic normality of the test statistics. Clearly, additional large sample tests can be obtained in those competing risk situations for which estimates of parameters are available and are known to be asymptotically normal.

REFERENCES

[1]Bernoulli, Daniel: Essai d'une nouvelle analyse de la mortalité causée par la petite verole, & des avantages de l'inoculation pour la prévenir. *Académie des Sciences, Paris. Historie avec les Mémoires,* 1760. pp. 1-45.

[2]Halley, E.: An estimate of the degrees of the mortality of mankind, drawn from curious tables of the births and funerals at the city of Breslau. *Philos. Trans. Roy. Soc.* (London) 17:569-610, 1693.

[3]Gompertz, B.: On the nature of the function expressive of the law of human mortality, *Philos. Trans. Roy. Soc.* (London) 115:513-583, 1825.

[4]Makeham, W. M.: On the law of mortality and the construction of annuity tables. *J. Inst. Actuaries* 8:1860.

[5]Makeham, W. M.: On an application of the theory of the composition of decremental forces. *J. Inst. Actuaries* 18:317-322, 1874.

[6]Spurgeon, E. F.: Life Contingencies, 3rd ed. London. *Cambridge University Press,* 1932.

[7]Bernstein, F., Birnbaum, Z. W., and Achs, S.: Is or is not cancer dependent on age? *Am. J. Cancer* 37:298-311, 1939.

[8]Cornfield, J.: The estimation of the probability of developing a disease in the presence of competing risks. *Am. J. Public Health* 47:601-607, 1957.

[9]Fix, E., and Neyman, J.: A simple stochastic model of recovery, relapse, death and loss of patients. *Hum. Biol.* 23:205-241, 1951.

[10]Chiang, C. L.: *Introduction to Stochastic Processes in Biostatistics.* New York. Wiley, 1968.

[11]Chiang, C. L.: Neyman's Contribution to the Theory of Competing Risks, a Fix-Neyman Mode. Long Term Care Report. No. 2. School of Public Health, University of California, Berkeley, 1974.

[12]Shepps, M. C., and Perrin, E. B.: Changes in birth rates as a function of contraceptive effectiveness: Some applications of a stochastic model. *Am. J. Pub. Health* 53:1031-1046, 1963.

[13]Perrin, E. B.: Uses of stochastic models in the evaluation of population policies. II. Extensions of the results by computer simulation. *Proc. Fifth Berkeley Symp. Math. Stat. Probab.* 4:137-146, 1967.

[14]Shepps, M. C., and Menken, J. A.: Mathematical Models of Conception and Birth. Chicago. University of Chicago Press, 1973.

[15]Tsiatis, A.: A nonidentifiability aspect of the problem of competing risks. *Proc. Nat. Acad. Sci. USA* 72:20-22, 1975.

[16]Berman, S. M.: Note on extreme values, competing risks and semi-Markov processes. *Ann. Math. Stat.* 34:1104-1106, 1963.

[17]Rose, D. M.: An Investigation of Dependent Competing Risks. Unpublished doctoral thesis, University of Washington, 1973.

[18]Peterson, A. V., Jr.: Bounds for a Joint Distribution Function With Fixed Sub-Distribution Functions; Application to Competing Risks. Tech. Rep. No. 7. Stanford University, Division of Biostatistics, 1975.

[19]Langberg, N., Proschan, F., and Quinzi, A. J.: Transformations yielding reliability models based on independent random variables: A survey, P. R. Krishnaiah, ed., in *Applications of Statistics.* Amsterdam. North Holland Publishing Co., 1977. pp. 323-337.

[20]Langberg, N., Proschan, F., and Quinzi, A. J.: Estimating dependent life lengths, with applications to the theory of competing risks. To be published.

[21]Langberg, N., Proschan, F., and Quinzi, A. J.: Converting dependent models into independent ones, preserving essential features. *Ann. Probab.* 6: 1978. In press.

[22]Kaplan, E. L., and Meier, P.: Nonparametric estimation from incomplete observations. *J. Am. Stat. Assoc.* 53:457-481, 1958.

[23]Breslow, N., and Crowley, J.: A large sample study of the life table and product limit estimates under random censorship. *Ann. Stat.* 2:437-453, 1974.

[24]Crowley, J.: Estimation of Relative Risk in Survival Studies. Tech. Rep. No. 423. Department of Statistics, University of Wisconsin, Madison, 1975.

[25]Aalen, O.: Nonparametric inference in connection with multiple decrement models. *Scand. J. Stat.* 3:15-27, 1976.

[26]Hoel, D. G.: A representation of mortality data by competing risks. *Biometrics* 28:475-488, 1972.

[27]David, H. A.: On Chiang's proportionality assumption in the theory of competing risks. *Biometrics* 26:336-339, 1970.

[28]Moeschberger, M. L., and David, H. A.: Life tests under competing causes of failure and the theory of competing risks. *Biometrics* 27:909-933, 1971.

[29]Herman, R. J., and Patell, R. K. N.: Maximum likelihood estimation for multi-risk model. *Technometrics* 13:385-396, 1971.

[30]Marshall, A. W., and Olkin, I.: A multivariate exponential distribution. *J. Am. Stat. Assoc.* 62:30-44, 1967.

[31]Lee, L., and Thompson, W. A., Jr.: Results on failure time and pattern for the series system, in Frank Proschan and R. J. Serfling, eds. *Reliability and Biometry*. Philadelphia. Society for Industrial and Applied Mathematics, 1974. pp. 291-302.

[32]Feigl, P., and Zelen, M.: Estimation of exponential survival probabilities with concomitant information. *Biometrics* 21:826-838, 1965.

[33]Zippin, C., and Armitage, P.: Use of concomitant variables and incomplete survival information in the estimation of an exponential survival parameter. *Biometrics* 22:665-672, 1966.

[34]Cox, D. R.: Regression models and life tables (with discussion). *J. Roy. Stat. Soc. B* 34:187-220, 1972.

[35]Kalbfleisch, F. O., and Prentice, R. L.: Marginal likelihood based on Cox's regression and life model. *Biometrika* 60:267-278, 1973.

[36]Lagakos, S. W.: A stochastic model for censored survival data in the presence of an auxiliary variable. *Biometrics* 32:551-559, 1976.

[37]Holford, T. R.: Life tables with concomitant information. *Biometrics* 32:587-597, 1976.

[38]Gehan, E. A.: A generalized Wilcoxon test for comparing arbitrarily single-censored samples. *Biometrika* 52:203-223, 1965.

[39]Gilbert, J. P.: Random Censorship. Unpublished doctoral thesis, University of Chicago, 1962.

[40]Gehan, E. A.: A generalized two-sample Wilcoxon test for doubly censored data. *Biometrika* 52:650-653, 1965.

[41]Efron, B.: The two-sample problem with censored data. *Proc. Fifth Berkeley Symp. Math. Stat. Probab.* 4:831-853, 1967.

[42]Breslow, N.: A generalized Kruskal-Wallis test for comparing K samples subject to unequal patterns of censorship. *Biometrika* 57:579-594, 1970.

[43]Mantel, N.: Ranking procedures for arbitrarily restricted observation. *Biometrics* 23:65-78, 1967.

[44]Brown, B. W., Hollander, M., and Korwar, R. M.: Nonparametric tests for censored data, with application to heart transplant studies, in Frank Proschan and R. J. Serfling, eds. *Reliability and Biometry*. Philadelphia, Society for Industrial and Applied Mathematics, 1974. pp. 327-354.

*U.S. GOVERNMENT PRINTING OFFICE : 1979 0-281-259/1

VITAL AND HEALTH STATISTICS Series

Series 1. Programs and Collection Procedures.—Reports which describe the general programs of the National Center for Health Statistics and its offices and divisions and data collection methods used and include definitions and other material necessary for understanding the data.

Series 2. Data Evaluation and Methods Research.—Studies of new statistical methodology including experimental tests of new survey methods, studies of vital statistics collection methods, new analytical techniques, objective evaluations of reliability of collected data, and contributions to statistical theory.

Series 3. Analytical Studies.—Reports presenting analytical or interpretive studies based on vital and health statistics, carrying the analysis further than the expository types of reports in the other series.

Series 4. Documents and Committee Reports.—Final reports of major committees concerned with vital and health statistics and documents such as recommended model vital registration laws and revised birth and death certificates.

Series 10. Data From the Health Interview Survey.—Statistics on illness, accidental injuries, disability, use of hospital, medical, dental, and other services, and other health-related topics, all based on data collected in a continuing national household interview survey.

Series 11. Data From the Health Examination Survey and the Health and Nutrition Examination Survey.—Data from direct examination, testing, and measurement of national samples of the civilian noninstitutionalized population provide the basis for two types of reports: (1) estimates of the medically defined prevalence of specific diseases in the United States and the distributions of the population with respect to physical, physiological, and psychological characteristics and (2) analysis of relationships among the various measurements without reference to an explicit finite universe of persons.

Series 12. Data From the Institutionalized Population Surveys.—Discontinued effective 1975. Future reports from these surveys will be in Series 13.

Series 13. Data on Health Resources Utilization.—Statistics on the utilization of health manpower and facilities providing long-term care, ambulatory care, hospital care, and family planning services.

Series 14. Data on Health Resources: Manpower and Facilities.—Statistics on the numbers, geographic distribution, and characteristics of health resources including physicians, dentists, nurses, other health occupations, hospitals, nursing homes, and outpatient facilities.

Series 20. Data on Mortality.—Various statistics on mortality other than as included in regular annual or monthly reports. Special analyses by cause of death, age, and other demographic variables; geographic and time series analyses; and statistics on characteristics of deaths not available from the vital records based on sample surveys of those records.

Series 21. Data on Natality, Marriage, and Divorce.—Various statistics on natality, marriage, and divorce other than as included in regular annual or monthly reports. Special analyses by demographic variables; geographic and time series analyses; studies of fertility; and statistics on characteristics of births not available from the vital records based on sample surveys of those records.

Series 22. Data From the National Mortality and Natality Surveys.—Discontinued effective 1975. Future reports from these sample surveys based on vital records will be included in Series 20 and 21, respectively.

Series 23. Data From the National Survey of Family Growth.—Statistics on fertility, family formation and dissolution, family planning, and related maternal and infant health topics derived from a biennial survey of a nationwide probability sample of ever-married women 15-44 years of age.

For a list of titles of reports published in these series, write to:

Scientific and Technical Information Branch
National Center for Health Statistics
Public Health Service
Hyattsville, Md. 20782